

EXISTENCE AND UNIQUENESS OF THE  $L^1$ -KARCHER MEAN

YONGDO LIM AND MIKLÓS PÁLFIA

ABSTRACT. We extend the domain of the Karcher mean  $\Lambda$  of positive operators on a Hilbert space to  $L^1$ -Borel probability measures on the cone of positive operators equipped with the Thompson part metric. We establish existence and uniqueness of  $\Lambda$  as the solution of the Karcher equation by developing a nonlinear ODE theory for the relative operator entropy in the spirit of Crandall-Liggett, such that the solutions of the Karcher equation are the stationary solutions of the ODE, and all solution curves enjoy the exponential contraction estimate. This is possible despite the facts that the Thompson metric is non-Euclidean, non-differentiable, non-commutative as a metric space as well as non-separable, and the positive cone is non-locally compact as a manifold. As further applications of the ODE approach, we prove the norm convergence conjecture of the power means of positive operators to the Karcher mean, and a Trotter-Kato product formula for the nonlinear semi-groups explicitly expressed by compositions of two-variable geometric means. This can be regarded as a nonlinear continuous-time version of the law of large numbers.

## 1. INTRODUCTION

Let  $\mathbb{S}$  denote the real vector space of self-adjoint operators,  $\mathbb{P} \subset \mathbb{S}$  denote the cone of positive definite operators on a Hilbert space  $\mathcal{H}$  equipped with the operator norm  $\|\cdot\|$ . Let

$$d_\infty(A, B) := \|\log(A^{-1/2}BA^{-1/2})\| = \text{spr}(\log(A^{-1}B))$$

denote the Thompson metric, which turns  $(\mathbb{P}, d_\infty)$  into a complete metric space such that the topology generated by  $d_\infty$  agrees with the relative operator norm topology [31]. The Karcher mean [12, 29], originally defined on  $\mathbb{P}$  for finite dimensional  $\mathcal{H}$  [3, 23] as a non-commutative generalization of the geometric mean [15], has been intensively investigated in the last decade [4, 19, 18]. For a  $k$ -tuple of operators  $\mathbb{A} := (A_1, \dots, A_k)$  with corresponding weight  $\omega := (w_1, \dots, w_k)$  where  $A_i \in \mathbb{P}$ ,  $w_i > 0$  and  $\sum_{i=1}^k w_i = 1$  the Karcher mean  $\Lambda(\omega, \mathbb{A})$  is defined as the unique solution of the Karcher equation

$$(1) \quad \sum_{i=1}^k w_i \log(X^{-1/2}A_iX^{-1/2}) = 0$$

for  $X \in \mathbb{P}$ . The existence and uniqueness in the infinite dimensional case was proved by Lawson-Lim [18], generalizing the approximation technique of power

---

*Date:* March 14, 2017.

*2000 Mathematics Subject Classification.* Primary 47A56, 47A64 Secondary 58B20.

*Key words and phrases.* operator mean, Karcher mean, relative operator entropy.

means given in the finite dimensional case by Lim-Pálfi [19]. The power mean  $P_t(\omega, \mathbb{A})$  for  $t \in (0, 1]$  is defined as the unique solution of the operator equation

$$(2) \quad \sum_{i=1}^k w_i X \#_t A_i = X$$

where  $A \#_t B = A^{1/2} (A^{-1/2} B A^{-1/2})^t A^{1/2}$  is the geometric mean of  $A, B \in \mathbb{P}$ . It is proved in [18] that  $t \mapsto P_t(\omega, \mathbb{A})$  is a decreasing sequence in the strong operator topology, strong operator converging to  $\Lambda(\omega, \mathbb{A})$ , extending the result of [19]. A further generalization of  $\Lambda$  to Borel probability measures with bounded support has been done in [14, 27] by integrating with respect to a Borel probability measure  $\mu$  in (1) and (2) instead of taking sums.

Let  $\mathcal{P}^1(\mathbb{P})$  denote the convex set of  $\tau$ -additive Borel probability measures  $\mu$  on  $(\mathbb{P}, \mathcal{B}(\mathbb{P}))$  such that  $\int_{\mathbb{P}} d_{\infty}(X, A) d\mu(A) < +\infty$  for all  $X \in \mathbb{P}$ . Recall that a Borel measure  $\mu$  is  $\tau$ -additive if  $\mu(\bigcup_{\alpha} U_{\alpha}) = \sup_{\alpha} \mu(U_{\alpha})$  for all directed families  $\{U_{\alpha} : \alpha \in D\}$  of open sets. In this paper for  $\mu \in \mathcal{P}^1(\mathbb{P})$  we consider the operator equation

$$(3) \quad \int_{\mathbb{P}} \log(X^{-1/2} A_i X^{-1/2}) d\mu(A) = 0,$$

and establish the existence and uniqueness of the solution for  $X \in \mathbb{P}$  which provides the extension of the map  $\Lambda(\cdot)$  to the case of  $L^1$ -probability measures over the infinite dimensional cone  $(\mathbb{P}, d_{\infty})$ . In particular existence is established by approximation with finitely supported measures in the  $L^1$ -Wasserstein distance, by extending the fundamental  $L^1$ -Wasserstein contraction property

$$d_{\infty}(\Lambda(\mu), \Lambda(\nu)) \leq W_1(\mu, \nu)$$

for any  $\mu, \nu \in \mathcal{P}^1(\mathbb{P})$  originally introduced by Sturm on CAT(0)-spaces [29]. To prove the uniqueness of solution of (3), we develop a nonlinear ODE theory for  $\Lambda$  by considering the Cauchy problem

$$(4) \quad \begin{aligned} X(0) &:= X \in \mathbb{P}, \\ \dot{X}(t) &= \int_{\mathbb{P}} \log_{X(t)} A d\mu(A), \end{aligned}$$

for  $t \in [0, \infty)$ , where  $\log_X A = X^{1/2} \log(X^{-1/2} A X^{-1/2}) X^{1/2}$  is the relative operator entropy [8, 9]. We prove that the solutions of (4) can be constructed by a discrete backward Euler-scheme converging in the  $d_{\infty}$  distance, generalizing the classical Crandall-Liggett techniques developed in Banach spaces [5]. In order to obtain the discretizations, we introduce the nonlinear resolvent

$$(5) \quad J_{\lambda}^{\mu}(X) := \Lambda\left(\frac{\lambda}{\lambda+1}\mu + \frac{1}{\lambda+1}\delta_X\right)$$

for  $\lambda > 0$ . The advantage is that  $J_{\lambda}^{\mu}$  is a strict contraction with respect to  $d_{\infty}$ , and satisfies the *resolvent identity* necessary to obtain the  $O(\sqrt{\lambda})$  convergence rate estimate of the exponential formula to the solution of (4) along with the semigroup property in [5]. We further obtain the exponential contraction rate estimate

$$d_{\infty}(\gamma(t), \eta(t)) \leq e^{-t} d_{\infty}(\gamma(0), \eta(0))$$

valid for two solution curves of (4) with varying initial points. This large time behavior will ensure the uniqueness of stationary points, thus the uniqueness of the solution of (3).

Furthermore, using the same exponential contraction estimate, we perform additional analysis of the Frechét-differential of the left hand side of (3) at the unique solution  $\Lambda(\mu)$ , eventually proving the norm convergence conjecture of power means  $P_t$  to  $\Lambda$  as  $t \rightarrow 0$ , a problem first mentioned in [18] as a possible strengthening of the strong operator convergence. From the analysis of the Frechét-differential, a resolvent convergence also follows which leads to a continuous-time Trotter-Kato-type product formula that is closely related to the law of large numbers of Sturm [29] and its deterministic "nodice" counterparts proved in [11, 20] valid in CAT(0)-spaces. In particular we prove that for a sequence  $\mu_n$  of finitely supported probability measures  $W_1$ -convergent to  $\mu$ , we have

$$\lim_{n \rightarrow \infty} S^{\mu_n}(t) = S^\mu(t)$$

in  $d_\infty$ , where  $S^\mu(t)$  denotes the solution of the Cauchy problem (4) corresponding to  $\mu$ . Under the assumption  $\mu_n = \sum_{i=1}^n \frac{1}{n} \delta_{Y_i}$ , we also prove the explicit product formula

$$\lim_{m \rightarrow \infty} \left( F_{t/m}^{\mu_n} \right)^m = S^{\mu_n}(t)$$

in  $d_\infty$ , where  $F_\rho^{\mu_n} := J_{\rho/n}^{\delta_{Y_n}} \circ \dots \circ J_{\rho/n}^{\delta_{Y_1}}$  with  $J_\rho^{\delta_A}(X) := X \#_{\frac{\rho}{\rho+1}} A$  in the spirit of (5). The above formula is advantageous, since it only contains iterated geometric means of only two operators, hence explicitly calculable.

It must be noted that although similar results are available for the Karcher mean in the CAT(0) [2, 28, 29] or even CAT( $k$ ) [25, 26] setting, all these techniques break down in the infinite dimensional case  $(\mathbb{P}, d_\infty)$  due to the nonexistence of a convex potential function so that the solutions of (4) is a gradient flow, the non-differentiability of the squared distance function  $d_\infty^2(A, \cdot)$  and the "non-commutativity" in the sense of [26] of the operator norm  $\|\cdot\|$  appearing in the formula for  $d_\infty$ . Even proving the resolvent convergence necessary for a Trotter-Kato product formula is non-trivial in the infinite dimensional case due to the lack of local compactness of the manifold  $\mathbb{P}$ .

The paper is organized as follows. In section 2 we gather all the necessary information available for the Karcher mean of finitely supported measures in relation with the distance  $d_\infty$  and the  $L^1$ -Wasserstein distance  $W_1$ , and in section 3 we extend the domain of  $\Lambda$  to  $\mathcal{P}^1(\mathbb{P})$  by  $W_1$ -continuity. In section 4 we develop the ODE theory corresponding to (4) by generalizing the argumentation in [5]. In section 5 we briefly gather the consequences of the exponential contraction estimate and in section 6 we develop the theory of approximation semigroups for the ODE flow (4). In section 7 we establish the  $d_\infty$  convergence of the approximating resolvents necessary for establishing the Trotter-Kato formula of section 6 by combining the large time behavior of the solutions of (4) with operator theoretic techniques. As a byproduct we prove the norm convergence conjecture of  $P_t$  to  $\Lambda$  as  $t \rightarrow 0$ . The last section gathers the consequences of the earlier sections, establishing a continuous-time result corresponding to the law of large numbers.

## 2. PRELIMINARIES

Let  $\mathcal{P}^1(\mathbb{P})$  denote the convex set of  $\tau$ -additive Borel probability measures  $\mu$  on  $(\mathbb{P}, \mathcal{B}(\mathbb{P}))$  such that  $\int_{\mathbb{P}} d_{\infty}(X, A) d\mu(A) < +\infty$  for all  $X \in \mathbb{P}$ . For  $\tau$ -additive measures one may realize the complement of the support as a directed union of open sets of measure 0 hence the complement has measure 0, and the support has measure 1. We consider  $\tau$ -additive measures, since the following is not true in general for  $\sigma$ -additive Borel probability measures, however it holds for  $\tau$ -additive ones:

**Proposition 2.1** (Theorem 8.3.2. & Example 8.1.6. [7]). *The topology generated by the Wasserstein metric  $W_1(\cdot, \cdot)$  on  $\mathcal{P}^1(\mathbb{P})$  agrees with the weak-\* (also called weak) topology of  $\mathcal{P}^1(\mathbb{P})$ . Moreover finitely supported probability measures are  $W_1$ -dense in  $\mathcal{P}^1(\mathbb{P})$ .*

Looks like the separability of support problem for  $\sigma$ -additive Borel probability measures is a deep question, with definitive answer depending on the properties of certain extensions of ZFC set theory, see Exercise 13 of chapter 6 in [13]. For the separability of the support of  $\tau$ -additive measures over metric spaces, see [16]. We provide here an alternative proof by approximation based on a result of [10].

**Proposition 2.2.** *Let  $\mu \in \mathcal{P}^1(\mathbb{P})$ . Then the support  $\text{supp}(\mu)$  is separable.*

*Proof.* By Proposition 2.1 we can approximate  $\mu$  by a sequence of finitely supported probability measures  $\mu_n \in \mathcal{P}^1(\mathbb{P})$  in the metric  $W_1$ . Then the main Theorem in [10] states that the weak-\* limit, hence  $W_1$ -limit in our case, of a convergent sequence of Borel probability measures with separable supports has also separable support. Thus the assertion follows.  $\square$

**Definition 2.1** (strong measurability, Bochner integral). Let  $(\Omega, \Sigma, \mu)$  be finite measure space and let  $f : \Omega \mapsto \mathbb{P}$ . Then  $f$  is *strongly measurable* if there exists a sequence of simple functions  $f_n$ , such that  $\lim_{n \rightarrow \infty} f_n(\omega) = f(\omega)$  almost surely.

The function  $f : \Omega \mapsto \mathbb{P}$  is *Bochner integrable* if the following are satisfied:

- (1)  $f$  is strongly measurable;
- (2) there exists a sequence of simple functions  $f_n$ , such that  $\lim_{n \rightarrow \infty} \int_{\Omega} \|f(\omega) - f_n(\omega)\| d\mu(\omega) = 0$

In this case we define the *Bochner integral* of  $f$  by

$$\int_{\Omega} f(\omega) d\mu(\omega) := \lim_{n \rightarrow \infty} \int_{\Omega} f_n(\omega) d\mu(\omega).$$

It is well known that a strongly measurable function  $f$  on a finite measure space  $(\Omega, \Sigma, \mu)$  is Bochner integrable if and only if  $\int_{\Omega} \|f(\omega)\| d\mu(\omega) < \infty$ .

The logarithm map  $\log : \mathbb{P} \rightarrow \mathbb{S}$  is differentiable and is contractive from the exponential metric increasing (EMI) property ([17])

$$(6) \quad \|\log X - \log Y\| \leq d_{\infty}(X, Y), \quad X, Y \in \mathbb{P}.$$

This property reflects the seminegative curvature of the Thompson metric, which can be realized as a Banach-Finsler metric arising from the Banach space norm on  $\mathbb{S}$ : for  $A \in \mathbb{P}$ , the Finsler norm of  $X \in T_A \mathbb{P} = \mathbb{S}$  is given by  $\|X\|_A = \|A^{-1/2} X A^{-1/2}\|$  and the exponential and logarithm maps are

$$(7) \quad \exp_A(X) = A^{1/2} \exp(A^{-1/2} X A^{-1/2}) A^{1/2},$$

$$(8) \quad \log_A(X) = A^{1/2} \log(A^{-1/2} X A^{-1/2}) A^{1/2}.$$

Notice that also  $\log_A X = A \log(A^{-1}X)$ .

**Lemma 2.3.** *For all  $\mu \in \mathcal{P}^1(\mathbb{P})$  and  $X \in \mathbb{P}$ , the Bochner integral  $\int_{\mathbb{P}} \log_X A d\mu(A)$  exists.*

*Proof.* First of all, notice that  $A \mapsto X \log(X^{-1}A)$  is strongly measurable. Indeed since  $A \mapsto X \log(X^{-1}A)$  is norm continuous, hence  $d_{\infty}$  to norm continuous and it is almost separably valued, thus by the Pettis measurability theorem it is strongly measurable. Then

$$\begin{aligned} \int_{\mathbb{P}} \|X \log(X^{-1}A)\| d\mu(A) &\leq \int_{\mathbb{P}} \|X^{1/2}\| \|\log(X^{-1/2}AX^{-1/2})\| \|X^{1/2}\| d\mu(A) \\ &= \|X\| \int_{\mathbb{P}} \|\log(X^{-1/2}AX^{-1/2})\| d\mu(A) \\ &= \|X\| \int_{\mathbb{P}} d_{\infty}(X, A) d\mu(A) < \infty \end{aligned}$$

which shows Bochner integrability.  $\square$

**Definition 2.2** (Karcher equation/mean). For a  $\mu \in \mathcal{P}^1(\mathbb{P})$  the *Karcher equation* is defined as

$$(9) \quad \int_{\mathbb{P}} \log_X A d\mu(A) = 0,$$

where  $X \in \mathbb{P}$ . If (9) has a unique solution in  $X \in \mathbb{P}$ , then it is called the Karcher mean and is denoted by  $\Lambda(\mu)$ .

**Definition 2.3** (Weighted geometric mean). Let  $A, B \in \mathbb{P}$  and  $t \in [0, 1]$ . Then for  $(1-t)\delta_A + t\delta_B =: \mu \in \mathcal{P}^1(\mathbb{P})$  the Karcher equation

$$\int_{\mathbb{P}} \log_X A d\mu(A) = (1-t) \log_X A + t \log_X B = 0$$

has a unique solution  $A \#_t B = \Lambda(\mu)$  called the *weighted geometric mean* and

$$A \#_t B = A^{1/2} \left( A^{-1/2} B A^{-1/2} \right)^t A^{1/2} = A (A^{-1} B)^t.$$

By the dominated convergence theorem and Lemma 2.3 we have the following:

**Lemma 2.4.** *For each  $X \in \mathbb{P}$  and  $\mu \in \mathcal{P}^1(\mathbb{P})$  the function  $X \mapsto \int_{\mathbb{P}} \log_X A d\mu(A)$  is  $d_{\infty}$  to norm continuous.*

*Proof.* Pick a sequence  $X_n \rightarrow X$  in the  $d_{\infty}$  topology in  $\mathbb{P}$ . Then

$$\begin{aligned} (10) \quad &\left\| \int_{\mathbb{P}} \log_{X_n} A d\mu(A) - \int_{\mathbb{P}} \log_X A d\mu(A) \right\| \\ &\leq \int_{\mathbb{P}} \|\log_{X_n} A - \log_X A\| d\mu(A) \\ &\leq \int_{\mathbb{P}} \|\log_{X_n} A\| + \|\log_X A\| d\mu(A) \\ &\leq \|X_n\| \int_{\mathbb{P}} d_{\infty}(X_n, A) d\mu(A) + \|X\| \int_{\mathbb{P}} d_{\infty}(X, A) d\mu(A) < \infty, \end{aligned}$$

thus  $\|\log_{X_n} A - \log_X A\|$  is integrable. Since  $d_{\infty}$  agrees with the relative norm topology, we have that

$$F_n(A) := \|\log_{X_n} A - \log_X A\| \rightarrow 0$$

point-wisely for every  $A \in \mathbb{P}$  as  $n \rightarrow \infty$ . Then by the dominated convergence theorem we obtain

$$\lim_{n \rightarrow \infty} \int_{\mathbb{P}} \|\log_{X_n} A - \log_X A\| d\mu(A) = \int_{\mathbb{P}} \lim_{n \rightarrow \infty} \|\log_{X_n} A - \log_X A\| d\mu(A) = 0.$$

In view of (10) this proves the assertion.  $\square$

For some further known facts below, see for example [18].

**Lemma 2.5.** *For a fixed  $A \in \mathbb{P}$  and  $t \in [0, 1]$  the function  $f(X) := A \#_t X$  is a contraction on  $(\mathbb{P}, d_\infty)$  with Lipschitz constant  $(1 - t)$ .*

**Proposition 2.6** (Power means, cf. [18, 19]). *Let  $t \in (0, 1]$ ,  $A_i \in \mathbb{P}$  for  $1 \leq i \leq n$  and let  $\omega = (w_1, \dots, w_n)$  be a probability vector so that  $\mu = \sum_{i=1}^n w_i \delta_{A_i} \in \mathcal{P}^1(\mathbb{P})$ . Then the function*

$$f(X) := \int_{\mathbb{P}} X \#_t A d\mu(A)$$

*is a contraction on  $(\mathbb{P}, d_\infty)$  with Lipschitz constant  $(1 - t)$ , and thus the operator equation*

$$(11) \quad \int_{\mathbb{P}} X \#_t A d\mu(A) = X$$

*has a unique solution in  $X \in \mathbb{P}$  which is denoted by  $P_t(\mu)$ .*

**Theorem 2.7** (see [18, 19]). *Let  $t \in (0, 1]$ ,  $A_i \in \mathbb{P}$  for  $1 \leq i \leq n$  and let  $\omega = (w_1, \dots, w_n)$  be a probability vector so that  $\mu = \sum_{i=1}^n w_i \delta_{A_i} \in \mathcal{P}^1(\mathbb{P})$ . Then for  $1 \geq s \geq t > 0$  we have  $P_s(\mu) \geq P_t(\mu)$  and the strong operator limit*

$$(12) \quad X_0 := \lim_{t \rightarrow 0+} P_t(\mu)$$

*exists and  $X_0 = \Lambda(\mu)$ .*

By the above we define  $P_0(\mu) := \Lambda(\mu)$ .

**Proposition 2.8** ([18, 19]). *The function  $P_t(\cdot)$  is operator monotone. That is, let  $t \in [0, 1]$ ,  $A_i \leq B_i \in \mathbb{P}$  for  $1 \leq i \leq n$  and let  $\omega = (w_1, \dots, w_n)$  be a probability vector so that  $\mu = \sum_{i=1}^n w_i \delta_{A_i}, \nu = \sum_{i=1}^n w_i \delta_{B_i} \in \mathcal{P}^1(\mathbb{P})$ . Then*

$$(13) \quad P_t(\mu) \leq P_t(\nu).$$

**Theorem 2.9** (see Theorem 6.4. [18]). *Let  $A_i \in \mathbb{P}$  for  $1 \leq i \leq n$  and let  $\omega = (w_1, \dots, w_n)$  be a probability vector. Then for  $\mu = \sum_{i=1}^n w_i \delta_{A_i}$  the equation (9) has a unique positive definite solution  $\Lambda(\mu)$ .*

*In the special case  $n = 2$ , we have*

$$(14) \quad \Lambda((1 - t)\delta_A + t\delta_B) = A \#_t B$$

*for any  $t \in [0, 1]$ ,  $A, B \in \mathbb{P}$ .*

**Proposition 2.10** (see Proposition 2.5. [18]). *Let  $A_i, B_i \in \mathbb{P}$  for  $1 \leq i \leq n$ . Then  $\Lambda$  for  $\mu = \frac{1}{n} \sum_{i=1}^n \delta_{A_i}$  and  $\nu = \frac{1}{n} \sum_{i=1}^n \delta_{B_i}$  satisfies*

$$(15) \quad d_\infty(\Lambda(\mu), \Lambda(\nu)) \leq \sum_{i=1}^n \frac{1}{n} d_\infty(A_i, B_i),$$

in particular by permutation invariance of  $\Lambda$  in the variables  $(A_1, \dots, A_n)$  we have

$$(16) \quad d_\infty(\Lambda(\mu), \Lambda(\nu)) \leq \min_{\sigma \in S_n} \sum_{i=1}^n \frac{1}{n} d_\infty(A_i, B_{\sigma(i)}) = W_1(\mu, \nu).$$

### 3. EXTENSION OF $\Lambda$ BY $W_1$ -CONTINUITY

We extend  $\Lambda$  and its contraction properties by using continuity and contraction property of it with respect to  $W_1$ , along with the approximation properties of  $\mathcal{P}^1(\mathbb{P})$  with respect to the metric  $W_1$ .

**Lemma 3.1.** *Let  $X, Y \in \mathbb{P}$  and  $\mu_n, \mu \in \mathcal{P}^1(\mathbb{P})$  and  $\text{supp}(\mu_n), \text{supp}(\mu) \subseteq Z \subset \mathbb{P}$  where  $Z$  is closed and separable. Assume also that  $X \rightarrow Y$  in  $d_\infty$ ,  $\mu_n \rightarrow \mu$  in  $W_1$ . Then*

$$\int_{\mathbb{P}} \log_X \text{Ad}\mu_n(A) \rightarrow \int_{\mathbb{P}} \log_Y \text{Ad}\mu(A)$$

in the weak Banach space topology.

*Proof.* Let  $x, y \in \mathbb{P}$  and  $\mu, \nu \in \mathcal{P}^1(\mathbb{P})$ . For any real-valued norm continuous linear functional  $l^*$  we have

$$(17) \quad \begin{aligned} & \left| \left\langle \int_{\mathbb{P}} \log_x \text{Ad}\mu_n(A) - \int_{\mathbb{P}} \log_y \text{Ad}\mu(A), l^* \right\rangle \right| \\ & \leq \left| \left\langle \int_{\mathbb{P}} \log_x \text{Ad}\mu(A) - \int_{\mathbb{P}} \log_y \text{Ad}\mu(A), l^* \right\rangle \right| \\ & \quad + \left| \left\langle \int_{\mathbb{P}} \log_x \text{Ad}\mu_n(A) - \int_{\mathbb{P}} \log_x \text{Ad}\mu(A), l^* \right\rangle \right| \\ & \leq \left| \left\langle \int_{\mathbb{P}} \log_x \text{Ad}\mu(A), l^* \right\rangle - \left\langle \int_{\mathbb{P}} \log_y \text{Ad}\mu(A), l^* \right\rangle \right| \\ & \quad + \left| \int_{\mathbb{P}} \langle \log_x A, l^* \rangle d\mu_n(A) - \int_{\mathbb{P}} \langle \log_x A, l^* \rangle d\mu(A) \right|. \end{aligned}$$

If  $x \rightarrow y$  in  $d_\infty$ , then the first term in the above converges to 0 by Lemma 2.4. Now the complete metric space  $(Z, d_\infty)$  is separable, so we can apply some well known theorems for the metric  $W_1$  restricted for probability measures with support included in  $Z$ . In fact Proposition 7.1.5 in [1] tells us that  $\mu_n$  has uniformly integrable 1-moments, i.e.

$$\lim_{R \rightarrow \infty} \sup_n \int_{d_\infty(x, A) \geq R} d_\infty(x, A) d\mu_n(A) = 0.$$

Now, for the second term in (17) we have the estimates

$$\begin{aligned} \int_{\mathbb{P}} |\langle \log_x A, l^* \rangle| d\mu_n(A) & \leq \|l^*\|_* \|x\| \int_{\mathbb{P}} \|\log(x^{-1/2} A x^{-1/2})\| d\mu_n(A) \\ & \leq \|l^*\|_* \|x\| \int_{\mathbb{P}} d_\infty(x, A) d\mu_n(A), \end{aligned}$$

which means that  $\int_{\mathbb{P}} |\langle \log_x A, l^* \rangle| d\mu_n(A)$  is uniformly integrable as well. In [7] Lemma 8.4.3. says that if  $\xi_\alpha \rightarrow \xi$  in the weak-\* topology for Baire probability measures  $\xi_\alpha, \xi$  on a topological space  $X$ , then for every real-valued continuous function  $f$  on  $X$  satisfying  $\lim_{R \rightarrow \infty} \sup_\alpha \int_{|f| \geq R} |f| d\xi_\alpha = 0$ , we have  $\lim_\alpha \int_X f d\xi_\alpha = \int_X f d\xi$ . Thus the second term of (17) also converges to 0.  $\square$

**Theorem 3.2.** *For all  $\mu \in \mathcal{P}^1(\mathbb{P})$  there exists a solution of (9) denoted by  $\Lambda(\mu)$  (with slight abuse of notation), which satisfies*

$$(18) \quad d_\infty(\Lambda(\mu), \Lambda(\nu)) \leq W_1(\mu, \nu)$$

for all  $\nu \in \mathcal{P}^1(\mathbb{P})$ .

*Proof.* Let  $\mu \in \mathcal{P}^1(\mathbb{P})$ . Then by Proposition 2.1 there exists a  $W_1$ -convergent sequence of finitely supported probability measures  $\mu_n \in \mathcal{P}^1(\mathbb{P})$  such that  $W_1(\mu, \mu_n) \rightarrow 0$ . By Theorem 2.9  $\Lambda(\mu_n)$  exists for any  $n$  in the index set. We also have that  $W_1(\mu_m, \mu_n) \rightarrow 0$  as  $m, n \rightarrow \infty$  and by (16) it follows that  $d_\infty(\Lambda(\mu_m), \Lambda(\mu_n)) \rightarrow 0$  as  $m, n \rightarrow \infty$ , i.e.  $\Lambda(\mu_n)$  is a  $d_\infty$  Cauchy sequence. Thus we define

$$(19) \quad \tilde{\Lambda}(\mu) := \lim_{n \rightarrow \infty} \Lambda(\mu_n).$$

Since (18) holds by Proposition 2.10 for finitely supported probability measures, we extend (18) to the whole of  $\mathcal{P}^1(\mathbb{P})$  by  $W_1$ -continuity, using the  $W_1$ -density of finitely supported probability measures in  $\mathcal{P}^1(\mathbb{P})$ .

Then by construction for all  $n$  we have

$$\int_{\mathbb{P}} \log_{\Lambda(\mu_n)} Ad\mu_n(A) = 0,$$

thus by Lemma 3.1 we have

$$\int_{\mathbb{P}} \log_{\Lambda(\mu_n)} Ad\mu_n(A) \rightarrow \int_{\mathbb{P}} \log_{\tilde{\Lambda}(\mu)} Ad\mu(A)$$

weakly, that is

$$\int_{\mathbb{P}} \log_{\tilde{\Lambda}(\mu)} Ad\mu(A) = 0.$$

□

**Definition 3.1** (Karcher mean). Given a  $\mu \in \mathcal{P}^1(\mathbb{P})$ , we define  $\Lambda(\mu)$  as the limit obtained in Theorem 3.2. Notice that the limit does not depend on the actual approximating sequence of measures due to (18).

#### 4. AN ODE FLOW OF $\Lambda$

The fundamental  $W_1$ -contraction property (18) enables us to develop an ODE flow theory for  $\Lambda$  that resembles the gradient flow theory of its potential function in the finite dimensional CAT(0)-space case, see [20, 26] and the monograph [2]. Given a CAT( $\kappa$ )-space  $(X, d)$ , the *Moreau-Yoshida* resolvent of a lower semicontinuous function  $f$  is defined as

$$J_\lambda(x) := \arg \min_{y \in X} f(y) + \frac{1}{2\lambda} d^2(x, y)$$

for  $\lambda > 0$ . Then the gradient flow  $S(t)$  semigroup of  $f$  is defined as

$$S(t)x_0 := \lim_{n \rightarrow \infty} (J_{t/n})^n x_0$$

for  $t \in [0, \infty)$  and starting point  $x_0 \in X$ , see [2]. However in the infinite dimensional case substituting  $d_\infty$  in place of  $d$  in the above formulas leads to many difficulties, in particular  $d_\infty^2$  is not convex, moreover  $d_\infty$  is not differentiable, since the operator norm  $\|\cdot\|$  is an  $L^\infty$ -type norm, hence not smooth. Also the potential function  $f$  is not known to exist in the infinite dimensional case of  $\mathbb{P}$ , since there exists no finite trace on  $\mathcal{B}(\mathcal{H})$  to be used to define any Riemannian metric on  $\mathbb{P}$ . However if



we use the formulation of the critical point gradient equation corresponding to the definition of  $J_\lambda$  above, we can obtain a reasonable ODE theory in our setting for  $\Lambda$ .

**Definition 4.1** (Resolvent operator). Given  $\mu \in \mathcal{P}^1(\mathbb{P})$  we define the resolvent operator for  $\lambda > 0$  and  $X \in \mathbb{P}$  as

$$(20) \quad J_\lambda^\mu(X) := \Lambda \left( \frac{\lambda}{\lambda+1} \mu + \frac{1}{\lambda+1} \delta_X \right),$$

a solution we obtained in Theorem 3.2 of the Karcher equation

$$\frac{\lambda}{\lambda+1} \int_{\mathbb{P}} \log_Z \text{Ad}\mu(A) + \frac{1}{\lambda+1} \log_Z(X) = 0$$

for  $Z \in \mathbb{P}$  according to Definition 3.1.

The resolvent operator exists for  $\lambda \in [0, \infty]$  and provides a continuous pass from  $X$  to  $\Lambda(\mu)$ . An alternative such operator is

$$\Lambda(X \#_t \mu), \quad t \in [0, 1]$$

where

$$X \#_t \mu = f_*(\mu), \quad f(A) := X \#_t A.$$

We readily obtain the following fundamental contraction property of the resolvent.

**Proposition 4.1** (Resolvent contraction). Given  $\mu \in \mathcal{P}^1(\mathbb{P})$ , for  $\lambda > 0$  and  $X, Y \in \mathbb{P}$  we have

$$(21) \quad d_\infty(J_\lambda^\mu(X), J_\lambda^\mu(Y)) \leq \frac{1}{1+\lambda} d_\infty(X, Y).$$

*Proof.* Let  $\mu_\alpha \in \mathcal{P}^1(\mathbb{P})$  be a net of finitely supported measures  $W_1$ -converging to  $\mu$  by Proposition 2.1. Then by the triangle inequality and Proposition 2.10 we get

$$\begin{aligned} & d_\infty(J_\lambda^\mu(X), J_\lambda^\mu(Y)) \\ & \leq d_\infty(J_\lambda^\mu(X), J_\lambda^{\mu_\alpha}(X)) + d_\infty(J_\lambda^{\mu_\alpha}(X), J_\lambda^{\mu_\alpha}(Y)) + d_\infty(J_\lambda^{\mu_\alpha}(Y), J_\lambda^\mu(Y)) \\ & \leq d_\infty(J_\lambda^\mu(X), J_\lambda^{\mu_\alpha}(X)) + \frac{1}{1+\lambda} d_\infty(X, Y) + d_\infty(J_\lambda^{\mu_\alpha}(Y), J_\lambda^\mu(Y)). \end{aligned}$$

Since  $d_\infty(J_\lambda^\mu(Z), J_\lambda^{\mu_\alpha}(Z)) \rightarrow 0$  as  $\alpha \rightarrow \infty$  by (18), taking the limit  $\alpha \rightarrow \infty$  in the above chain of inequalities yields the assertion.  $\square$

**Proposition 4.2** (Resolvent identity). Given  $\mu \in \mathcal{P}^1(\mathbb{P})$ , for  $\tau > \lambda > 0$  and  $X \in \mathbb{P}$  we have

$$(22) \quad J_\tau^\mu(X) = J_\lambda^\mu \left( J_\tau^\mu(X) \#_{\frac{\lambda}{\tau}} X \right).$$

*Proof.* First suppose that  $\mu = \sum_{i=1}^n w_i \delta_{A_i}$  where  $A_i \in \mathbb{P}$  for  $1 \leq i \leq n$  and  $\omega = (w_1, \dots, w_n)$  a probability vector. By (20) we have

$$\tau \int_{\mathbb{P}} \log_{J_\tau^\mu(X)} \text{Ad}\mu(A) + \log_{J_\tau^\mu(X)} X = 0$$

and from that it follows that

$$\begin{aligned} & \lambda \int_{\mathbb{P}} \log_{J_\tau^\mu(X)} \text{Ad}\mu(A) + \frac{\lambda}{\tau} \log_{J_\tau^\mu(X)} X = 0, \\ & \lambda \int_{\mathbb{P}} \log_{J_\tau^\mu(X)} \text{Ad}\mu(A) + \log_{J_\tau^\mu(X)} \left( J_\tau^\mu(X) \#_{\frac{\lambda}{\tau}} X \right) = 0, \end{aligned}$$

and the above equation still uniquely determines  $J_\tau^\mu(X)$  as its only positive solution by Theorem 2.9, thus establishing (22) for finitely supported measures  $\mu$ .

The general  $\mu \in \mathcal{P}^1(\mathbb{P})$  case of (22) is obtained by approximating  $\mu$  in  $W_1$  by a net of finitely supported measures  $\mu_\alpha \in \mathcal{P}^1(\mathbb{P})$  and using (18) to show that  $J_\lambda^{\mu_\alpha}(X) \rightarrow J_\lambda^\mu(X)$  in  $d_\infty$  and also the fact that  $\#_t$  appearing in (22) is also  $d_\infty$ -continuous, hence obtaining (22) in the limit as  $\mu_\alpha \rightarrow \mu$  in  $W_1$ .  $\square$

**Proposition 4.3.** *Given  $\mu \in \mathcal{P}^1(\mathbb{P})$ ,  $\lambda > 0$  and  $X \in \mathbb{P}$  we have*

$$(23) \quad \begin{aligned} d_\infty(J_\lambda^\mu(X), X) &\leq \frac{\lambda}{1+\lambda} \int_{\mathbb{P}} d_\infty(X, A) d\mu(A) \\ d_\infty((J_\lambda^\mu)^n(X), X) &\leq n \frac{\lambda}{1+\lambda} \int_{\mathbb{P}} d_\infty(X, A) d\mu(A). \end{aligned}$$

*Proof.* By Theorem 3.2  $J_\lambda^\mu(X)$  is a solution of

$$(24) \quad \lambda \int_{\mathbb{P}} \log_{J_\lambda^\mu(X)} A d\mu(A) + \log_{J_\lambda^\mu(X)} X = 0,$$

hence we have

$$\begin{aligned} d_\infty(J_\lambda^\mu(X), X) &= \left\| \log \left( J_\lambda^\mu(X)^{-1/2} X J_\lambda^\mu(X)^{-1/2} \right) \right\| \\ &= \lambda \left\| \int_{\mathbb{P}} \log \left( J_\lambda^\mu(X)^{-1/2} A J_\lambda^\mu(X)^{-1/2} \right) d\mu(A) \right\| \\ &\leq \lambda \int_{\mathbb{P}} \left\| \log \left( J_\lambda^\mu(X)^{-1/2} A J_\lambda^\mu(X)^{-1/2} \right) \right\| d\mu(A) \\ &= \lambda \int_{\mathbb{P}} d_\infty(J_\lambda^\mu(X), A) d\mu(A) \end{aligned}$$

Given  $J_\lambda^\mu(X) \in \mathbb{P}$  we can solve (24) for  $X \in \mathbb{P}$ , thus by Proposition 4.1 we also have

$$d_\infty(J_\tau^\mu(X), X) = d_\infty \left( J_\tau^\mu(X), J_\tau^\mu \left( (J_\tau^\mu)^{-1}(X) \right) \right) \leq \frac{1}{1+\lambda} d_\infty \left( X, (J_\tau^\mu)^{-1}(X) \right),$$

hence the first inequality in (23) follows.

The second inequality in (23) follows from the first by the estimate

$$\begin{aligned} d_\infty((J_\lambda^\mu)^n(X), X) &\leq \sum_{i=0}^{n-1} d_\infty \left( (J_\lambda^\mu)^{n-i}(X), (J_\lambda^\mu)^{n-(i+1)}(X) \right) \\ &\leq \sum_{i=0}^{n-1} (1+\lambda)^{-n+(i+1)} d_\infty(J_\lambda^\mu(X), X) \\ &\leq n d_\infty(J_\lambda^\mu(X), X). \end{aligned}$$

$\square$

In what follows we will closely follow the arguments in [5] to construct the semigroups corresponding to the resolvent above.  $B(k, l)$  denotes the binomial coefficient.

**Lemma 4.4** (a variant of Lemma 1.3 cf. [5]). *Let  $\mu \in \mathcal{P}^1(\mathbb{P})$ ,  $\tau \geq \lambda > 0$ ;  $n \geq m$  be positive integers and  $X \in \mathbb{P}$ . Then*

$$\begin{aligned} & d_\infty \left( (J_\tau^\mu)^n(X), (J_\lambda^\mu)^m(X) \right) \\ & \leq (1 + \lambda)^{-n} \sum_{j=0}^{m-1} \alpha^j \beta^{n-j} B(n, j) d_\infty \left( (J_\tau^\mu)^{m-j}(X), X \right) \\ & \quad + \sum_{j=m}^n (1 + \lambda)^{-j} \alpha^m \beta^{j-m} B(j-1, m-1) d_\infty \left( (J_\lambda^\mu)^{n-j}(X), X \right) \end{aligned}$$

where  $\alpha = \frac{\lambda}{\tau}$  and  $\beta = \frac{\tau-\lambda}{\tau}$ .

*Proof.* For integers  $j$  and  $k$  satisfying  $0 \leq j \leq n$  and  $0 \leq k \leq m$ , put

$$a_{k,j} := d_\infty \left( (J_\lambda^\mu)^j(X), (J_\tau^\mu)^k(X) \right).$$

For  $j, k > 0$  by Proposition 4.1 and Proposition 4.2 we have

$$\begin{aligned} a_{k,j} &= d_\infty \left( (J_\lambda^\mu)^j(X), J_\lambda^\mu \left( (J_\tau^\mu)^k(X) \#_{\frac{\lambda}{\tau}} (J_\tau^\mu)^{k-1}(X) \right) \right) \\ &\leq (1 + \lambda)^{-1} d_\infty \left( (J_\lambda^\mu)^{j-1}(X), (J_\tau^\mu)^k(X) \#_{\frac{\lambda}{\tau}} (J_\tau^\mu)^{k-1}(X) \right) \\ &\leq (1 + \lambda)^{-1} \left[ \frac{\tau - \lambda}{\tau} d_\infty \left( (J_\lambda^\mu)^{j-1}(X), (J_\tau^\mu)^k(X) \right) \right. \\ &\quad \left. + \frac{\lambda}{\tau} d_\infty \left( (J_\lambda^\mu)^{j-1}(X), (J_\tau^\mu)^{k-1}(X) \right) \right] \\ &= (1 + \lambda)^{-1} \frac{\lambda}{\tau} a_{k-1,j-1} + (1 + \lambda)^{-1} \frac{\tau - \lambda}{\tau} a_{k,j-1}, \end{aligned}$$

where to obtain the second inequality we used Proposition 2.10 for  $\#_t$ . From here, the rest of the proof follows along the lines of Lemma 1.3 in [5].  $\square$

We quote the following Lemma 1.4. from [5]:

**Lemma 4.5.** *Let  $n \geq m > 0$  be integers, and  $\alpha, \beta$  positive numbers satisfying  $\alpha + \beta = 1$ . Then*

$$\sum_{j=0}^m B(n, j) \alpha^j \beta^{n-j} (m-j) \leq \sqrt{(n\alpha - m)^2 + n\alpha\beta},$$

and

$$\sum_{j=m}^n B(j-1, m-1) \alpha^m \beta^{j-m} (n-j) \leq \sqrt{\frac{m\beta}{\alpha^2} + \left( \frac{m\beta}{\alpha^2} + m - n \right)^2}.$$

**Theorem 4.6.** *For any  $X, Y \in \mathbb{P}$  and  $t > 0$  the curve*

$$(25) \quad S(t)X := \lim_{n \rightarrow \infty} \left( J_{t/n}^\mu \right)^n(X)$$

*exists where the limit is in the  $d_\infty$ -topology and it is Lipschitz-continuous on compact time intervals  $[0, T]$  for any  $T > 0$ . Moreover it satisfies the contraction property*

$$(26) \quad d_\infty(S(t)X, S(t)Y) \leq e^{-t} d_\infty(X, Y),$$

and for  $s > 0$  verifies the semigroup property

$$(27) \quad S(t+s)X = S(t)(S(s)X),$$

and the flow operator  $S : \mathbb{P} \times (0, \infty) \mapsto \mathbb{P}$  extends by  $d_\infty$ -continuity to  $S : \mathbb{P} \times [0, \infty) \mapsto \mathbb{P}$ .

*Proof.* The proof closely follows that of Theorem I in [5] using the previous estimates of this section. In particular for  $n \geq m > 0$  one obtains

$$(28) \quad d_\infty \left( \left( J_{t/n}^\mu \right)^n (X), \left( J_{t/m}^\mu \right)^m (X) \right) \leq 2t \left( \frac{1}{m} - \frac{1}{n} \right)^{1/2} \int_{\mathbb{P}} d_\infty(X, A) d\mu(A)$$

so  $\lim_{n \rightarrow \infty} \left( J_{t/n}^\mu \right)^n (X)$  exists proving (25). Also  $\left( J_{t/n}^\mu \right)^n$  satisfies

$$d_\infty \left( \left( J_{t/n}^\mu \right)^n (X), \left( J_{t/n}^\mu \right)^n (Y) \right) \leq \left( 1 + \frac{t}{n} \right)^{-n} d_\infty(X, Y),$$

hence also (26). We also have

$$(29) \quad d_\infty(S(s)X, S(t)X) \leq 2|s - t| \int_{\mathbb{P}} d_\infty(X, A) d\mu(A)$$

proving Lipschitz-continuity in  $t$  on compact time intervals. The proof of the semigroup property is exactly the same as in [5].  $\square$

Before stating the next result we need another auxiliary lemma describing the asymptotic behavior of  $J_{t/n}^\mu(X)$ .

**Lemma 4.7.** *Let  $\mu \in \mathcal{P}^1(\mathbb{P})$ ,  $\lambda > 0$  and  $X \in \mathbb{P}$ . Then*

$$(30) \quad \log_{J_\lambda^\mu(X)} X = X - J_\lambda^\mu(X) + O(\lambda^2).$$

*Proof.* Let  $C := \int_{\mathbb{P}} d_\infty(X, A) d\mu(A)$ . Then by Proposition 4.3

$$e^{-\lambda(1+\lambda)^{-1}C} - I \leq J_\lambda^\mu(X)^{-1/2} X J_\lambda^\mu(X)^{-1/2} - I \leq e^{\lambda(1+\lambda)^{-1}C} - I,$$

hence

$$e^{-\lambda C} - I \leq J_\lambda^\mu(X)^{-1/2} X J_\lambda^\mu(X)^{-1/2} - I \leq e^{\lambda C} - I,$$

which yields

$$(31) \quad \sum_{k=1}^{\infty} (-1)^k \frac{(\lambda C)^k}{k!} \leq J_\lambda^\mu(X)^{-1/2} X J_\lambda^\mu(X)^{-1/2} - I \leq \sum_{k=1}^{\infty} \frac{(\lambda C)^k}{k!}.$$

In view of the series expansion

$$\log(z) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{(z - I)^k}{k}$$

uniformly convergent for  $\|z - I\| < 1$ , we get

$$\begin{aligned} \log \left( J_\lambda^\mu(X)^{-1/2} X J_\lambda^\mu(X)^{-1/2} \right) \\ = J_\lambda^\mu(X)^{-1/2} X J_\lambda^\mu(X)^{-1/2} - I + O(\lambda^2), \end{aligned}$$

from which the assertion follows.  $\square$

The proof of the following theorem in essence is analogous to that of Theorem II in [5].

**Theorem 4.8.** *Let  $\mu \in \mathcal{P}^1(\mathbb{P})$  and  $X \in \mathbb{P}$ . Then for  $t > 0$ , the curve  $X(t) := S(t)X$  provides a strong solution of the Cauchy problem*

$$\begin{aligned} X(0) &:= X, \\ \dot{X}(t) &= \int_{\mathbb{P}} \log_{X(t)} \text{Ad}\mu(A), \end{aligned}$$

where the derivative  $\dot{X}(t)$  is the Fréchet-derivative.

*Proof.* Due to the semigroup property of  $S(t)$ , it is enough to check that

$$\lim_{t \rightarrow 0+} \frac{S(t)X - X}{t} = \int_{\mathbb{P}} \log_X \text{Ad}\mu(A)$$

where the limit is in the norm topology. We have

$$\begin{aligned} \frac{S(t)X - X}{t} &= \lim_{n \rightarrow \infty} \frac{\left(J_{t/n}^\mu\right)^n(X) - X}{t} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \frac{\sum_{i=0}^{n-1} J_{t/n}^\mu \left( \left(J_{t/n}^\mu\right)^i(X) \right) - \left(J_{t/n}^\mu\right)^i(X)}{t/n} \end{aligned}$$

and also

$$\frac{t}{n} \int_{\mathbb{P}} \log_{\left(J_{t/n}^\mu\right)^i(X)} \text{Ad}\mu(A) + \log_{\left(J_{t/n}^\mu\right)^i(X)} \left(J_{t/n}^\mu\right)^{i-1}(X) = 0.$$

Then by Lemma 4.7 we have

$$\frac{S(t)X - X}{t} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \int_{\mathbb{P}} \log_{\left(J_{t/n}^\mu\right)^i(X)} \text{Ad}\mu(A) + O\left(\frac{t}{n}\right),$$

which combined with the estimates in Proposition 4.3 and Lemma 2.4 proves the assertion.  $\square$

## 5. CONSEQUENCES FOR $\Lambda$

**Theorem 5.1.** *Assume  $\mu \in \mathcal{P}^1(\mathbb{P})$ . Then the Karcher equation (9) has a unique solution in  $\mathbb{P}$ .*

*Proof.* By Theorem 3.2 the Karcher equation (9) has a solution  $\Lambda(\mu)$  in  $\mathbb{P}$ . Assume that  $\tilde{\Lambda}(\mu) \in \mathbb{P}$  is another solution of (9) such that  $\tilde{\Lambda}(\mu) \neq \Lambda(\mu)$ . For  $t > 0$  consider the curves  $\gamma(t) := S(t)\Lambda(\mu)$  and  $\eta(t) := S(t)\tilde{\Lambda}(\mu)$ . Both curves solve the ODE in Theorem 4.8, but since both  $\Lambda(\mu)$  and  $\tilde{\Lambda}(\mu)$  solve the equation (9), the right hand side of the ODE in Theorem 4.8 vanishes for all time  $t > 0$ , hence the flows  $\gamma$  and  $\eta$  are both stationary,  $\gamma(t) = \Lambda(\mu)$  and  $\eta(t) = \tilde{\Lambda}(\mu)$ , thus  $d_\infty(\gamma(t), \eta(t)) = d_\infty(\Lambda(\mu), \tilde{\Lambda}(\mu)) > 0$  for almost all  $t > 0$ . However by (26) in Theorem 4.6 we have the strict contraction estimate

$$d_\infty(\gamma(t), \eta(t)) \leq e^{-t} d_\infty(\Lambda(\mu), \tilde{\Lambda}(\mu))$$

which for  $t > 0$  contradicts  $d_\infty(\Lambda(\mu), \tilde{\Lambda}(\mu)) > 0$ , hence we have  $d_\infty(\Lambda(\mu), \tilde{\Lambda}(\mu)) = 0$ .  $\square$

**Remark 5.1.** The proof of the Resolvent Identity Proposition 4.2 through the Karcher equation requires the uniqueness of  $\Lambda(\mu)$  as a solution of the Karcher equation for finitely supported measures, hence the uniqueness argument in [18]. It would be nice to get rid of this uniqueness assumption, so that we obtain the uniqueness of the solution of the Karcher equation directly by the above ODE flow theory alone.

**Remark 5.2.** Many properties of  $\Lambda$  now carries over to the  $L^1$ -setting. The interested reader can consult section 6 in [27] and [16]. In particular the stochastic order introduced in [14, 16] extends the usual element-wise order of uniformly finitely supported measures by introducing upper sets:  $U \subseteq \mathbb{P}$  is upper if for an  $X \in \mathbb{P}$  there exists an  $Y \in U$  such that  $Y \leq X$ , then  $X \in U$ . Then the *stochastic order* for  $\mu, \nu \in \mathcal{P}^1(\mathbb{P})$  is defined as  $\mu \leq \nu$  if  $\mu(U) \leq \nu(U)$  for all upper sets  $U \subseteq \mathbb{P}$ . Then the results in [16] applies and if  $\mu \leq \nu$  then  $\Lambda(\mu) \leq \Lambda(\nu)$ . This can also be proved by applying the results of section 6 in [14] to the infinite dimensional setting with the monotonicity results of [27] for measures with bounded support.

## 6. APPROXIMATING SEMIGROUPS AND TROTTER-KATO PRODUCT FORMULA

In this section we develop the theory of approximating semigroups that will lead to a Trotter-Kato product formula for the nonlinear ODE semigroups of the Karcher mean.

**Lemma 6.1.** *Let  $F : \mathbb{P} \mapsto \mathbb{P}$  be a nonexpansive map with respect to  $d_\infty$ . Let  $\lambda, \rho > 0$  and  $Y \in \mathbb{P}$ . Then the map*

$$G_{\lambda, \rho, Y}(X) := \Lambda \left( \frac{1}{1 + \lambda/\rho} \delta_Y + \frac{\lambda/\rho}{1 + \lambda/\rho} \delta_{F(X)} \right)$$

*is a strict contraction with Lipschitz constant  $\frac{\lambda/\rho}{1 + \lambda/\rho} < 1$ . Consequently the map  $G_{\lambda, \rho, Y}$  has a unique fixed point denoted by  $J_{\lambda, \rho}(Y)$ .*

*Proof.* By Proposition 2.10 for  $X_1, X_2 \in \mathbb{P}$  we get

$$d_\infty(G_{\lambda, \rho, Y}(X_1), G_{\lambda, \rho, Y}(X_2)) \leq \frac{\lambda/\rho}{1 + \lambda/\rho} d_\infty(F(X_1), F(X_2)) \leq \frac{\lambda/\rho}{1 + \lambda/\rho} d_\infty(X_1, X_2),$$

thus by Banach's fixed point theorem  $G_{\lambda, \rho, Y}$  has a unique fixed point denoted by  $J_{\lambda, \rho}(Y)$ , hence

$$(32) \quad J_{\lambda, \rho}(Y) = \Lambda \left( \frac{1}{1 + \lambda/\rho} \delta_Y + \frac{\lambda/\rho}{1 + \lambda/\rho} \delta_{F(J_{\lambda, \rho}(Y))} \right).$$

□

**Lemma 6.2.** *Let  $F : \mathbb{P} \mapsto \mathbb{P}$  be a nonexpansive map with respect to  $d_\infty$ . Then for  $\lambda, \rho > 0$  the map  $J_{\lambda, \rho} : \mathbb{P} \mapsto \mathbb{P}$  is nonexpansive.*

*Proof.* By Proposition 2.10 and Lemma 6.1 for  $X_1, X_2 \in \mathbb{P}$  and  $t := \frac{\lambda/\rho}{1 + \lambda/\rho} < 1$  we get

$$\begin{aligned} d_\infty(J_{\lambda, \rho}(X_1), J_{\lambda, \rho}(X_2)) &\leq (1 - t) d_\infty(X_1, X_2) + t d_\infty(F(J_{\lambda, \rho}(X_1)), F(J_{\lambda, \rho}(X_2))) \\ &\leq (1 - t) d_\infty(X_1, X_2) + t d_\infty(J_{\lambda, \rho}(X_1), J_{\lambda, \rho}(X_2)), \end{aligned}$$

from which  $d_\infty(J_{\lambda, \rho}(X_1), J_{\lambda, \rho}(X_2)) \leq d_\infty(X_1, X_2)$  follows. □

**Lemma 6.3.** *Let  $F : \mathbb{P} \mapsto \mathbb{P}$  be a nonexpansive map. Then for  $\lambda, \rho > 0$  and  $X \in \mathbb{P}$  we have*

$$(33) \quad \frac{d_\infty(X, J_{\lambda, \rho}(X))}{\lambda} \leq \frac{d_\infty(X, F(X))}{\rho}.$$

*Proof.* By Lemma 6.1 we have  $\lim_{n \rightarrow \infty} G_{\lambda, \rho, X}^n(X) = J_{\lambda, \rho}(X)$ . Then

$$\begin{aligned} d_\infty(X, J_{\lambda, \rho}(X)) &\leq \sum_{n=1}^{\infty} d_\infty(G_{\lambda, \rho, X}^{n-1}(X), G_{\lambda, \rho, X}^n(X)) \\ &\leq \sum_{n=1}^{\infty} \left( \frac{\lambda/\rho}{1 + \lambda/\rho} \right)^{n-1} d_\infty(X, G_{\lambda, \rho, X}(X)) \\ &\leq \frac{1}{1 - \frac{\lambda/\rho}{1 + \lambda/\rho}} d_\infty(X, G_{\lambda, \rho, X}(X)) = \frac{\lambda}{\rho} d_\infty(X, F(X)), \end{aligned}$$

since  $d_\infty(X, G_{\lambda, \rho, X}(X)) = \frac{\lambda/\rho}{1 + \lambda/\rho} d_\infty(X, F(X))$ .  $\square$

**Lemma 6.4** (Resolvent Identity). *Let  $F : \mathbb{P} \mapsto \mathbb{P}$  be a nonexpansive map. Then for  $\lambda > \mu, \rho > 0$  and  $X \in \mathbb{P}$  we have*

$$(34) \quad J_{\lambda, \rho}(X) = J_{\mu, \rho} \left( J_{\lambda, \rho}(X) \#_{\frac{\mu}{\lambda}} X \right).$$

*Proof.* First of all notice that for  $A, B \in \mathbb{P}$  we have that the curve  $c(t) = A \#_t B$  for  $t \in [0, 1]$  has the property that for any  $s \leq u \in [0, 1]$  the curve  $v(t) := c(s) \#_t c(u)$  is a connected subset of the curve  $c$ , i.e.  $v(t) = c(s + t(u - s))$ .

Now consider the curve  $\gamma(t) := X \#_t F(J_{\lambda, \rho}(X))$  for  $t \in [0, 1]$ . Then by definition it follows that the points  $J_{\lambda, \rho}(X)$  and  $Z := J_{\lambda, \rho}(X) \#_{\frac{\mu}{\lambda}} X$  are also on the curve  $\gamma$ , hence by the above  $c(t) := Z \#_t F(J_{\lambda, \rho}(X))$  for  $t \in [0, 1]$  is a connected subset of the curve  $\gamma$ . Then to conclude our assertion, by (32) and Theorem 2.9, it suffices to show that

$$\frac{d_\infty(Z, J_{\lambda, \rho}(X))}{d_\infty(Z, F(J_{\lambda, \rho}(X)))} = \frac{\mu}{\rho + \mu}.$$

Indeed, let  $a := d_\infty(X, F(J_{\lambda, \rho}(X)))$ ,  $b := d_\infty(J_{\lambda, \rho}(X), F(J_{\lambda, \rho}(X)))$  and  $d := d_\infty(Z, J_{\lambda, \rho}(X))$ . Then  $b = \frac{\rho}{\rho + \lambda} a$ ,  $d = (a - b) \frac{\mu}{\lambda} = \frac{\lambda}{\rho + \lambda} \frac{\mu}{\lambda} a = \frac{\mu}{\rho + \lambda} a$ , thus we have

$$\frac{d_\infty(Z, J_{\lambda, \rho}(X))}{d_\infty(Z, F(J_{\lambda, \rho}(X)))} = \frac{d}{d + b} = \frac{\mu}{\rho + \mu}.$$

$\square$

**Lemma 6.5.** *Let  $F : \mathbb{P} \mapsto \mathbb{P}$  be a nonexpansive map. Then for  $\lambda, \rho > 0, n \in \mathbb{N}$  and  $X \in \mathbb{P}$  we have*

$$(35) \quad d_\infty(J_{\lambda, \rho}^n(X), X) \leq n \frac{\lambda}{\rho} d_\infty(X, F(X)).$$

*Proof.* By the triangle inequality, Lemma 6.2 and Lemma 6.3 we have

$$\begin{aligned} d_\infty(J_{\lambda, \rho}^n(X), X) &\leq \sum_{i=1}^n d_\infty(J_{\lambda, \rho}^i(X), J_{\lambda, \rho}^{i-1}(X)) \\ &\leq n d_\infty(J_{\lambda, \rho}(X), X) \\ &\leq n \frac{\lambda}{\rho} d_\infty(X, F(X)). \end{aligned}$$

□

**Lemma 6.6.** *Let  $F : \mathbb{P} \mapsto \mathbb{P}$  be a nonexpansive map,  $\tau \geq \lambda > 0$ ,  $\rho > 0$ ;  $n \geq m$  be positive integers and  $X \in \mathbb{P}$ . Then*

$$\begin{aligned} & d_{\infty}((J_{\tau,\rho})^n(X), (J_{\lambda,\rho})^m(X)) \\ & \leq \sum_{j=0}^{m-1} \alpha^j \beta^{n-j} B(n, j) d_{\infty}((J_{\tau,\rho})^{m-j}(X), X) \\ & \quad + \sum_{j=m}^n \alpha^m \beta^{j-m} B(j-1, m-1) d_{\infty}((J_{\lambda,\rho})^{n-j}(X), X) \end{aligned}$$

where  $\alpha = \frac{\lambda}{\tau}$  and  $\beta = \frac{\tau-\lambda}{\tau}$ .

*Proof.* Using the Resolvent Identity Lemma 6.4, the estimates are obtained in the same way as in Lemma 4.4. □

**Theorem 6.7.** *Let  $F : \mathbb{P} \mapsto \mathbb{P}$  be a nonexpansive map. Then for any  $X, Y \in \mathbb{P}$  and  $t, \rho > 0$  the curve*

$$(36) \quad S_{\rho}(t)X := \lim_{n \rightarrow \infty} (J_{t/n, \rho})^n(X)$$

*exists where the limit is in the  $d_{\infty}$ -topology with estimate*

$$(37) \quad d_{\infty}(S_{\rho}(t)X, (J_{t/n, \rho})^n(X)) \leq \frac{2t}{\sqrt{n}} \frac{d_{\infty}(X, F(X))}{\rho},$$

*and satisfies the Lipschitz estimate*

$$(38) \quad d_{\infty}(S_{\rho}(t)X, S_{\rho}(s)X) \leq 2 \frac{d_{\infty}(X, F(X))}{\rho} |t - s|$$

*for any  $t, s \geq 0$ . Moreover it also satisfies the contraction property*

$$(39) \quad d_{\infty}(S_{\rho}(t)X, S_{\rho}(t)Y) \leq d_{\infty}(X, Y),$$

*for  $s > 0$  verifies the semigroup property*

$$(40) \quad S_{\rho}(t+s)X = S_{\rho}(t)(S_{\rho}(s)X),$$

*and the flow operator  $S_{\rho} : \mathbb{P} \times (0, \infty) \mapsto \mathbb{P}$  extends by  $d_{\infty}$ -continuity to  $S_{\rho} : \mathbb{P} \times [0, \infty) \mapsto \mathbb{P}$ .*



*Proof.* We closely follow the proof of Theorem 4.6. Using the previous lemmas we similarly obtain

$$\begin{aligned}
& d_\infty((J_{\tau,\rho})^n(X), (J_{\lambda,\rho})^m(X)) \\
& \leq \sum_{j=0}^{m-1} \alpha^j \beta^{n-j} B(n, j) d_\infty((J_{\tau,\rho})^{m-j}(X), X) \\
& \quad + \sum_{j=m}^n \alpha^m \beta^{j-m} B(j-1, m-1) d_\infty((J_{\lambda,\rho})^{n-j}(X), X) \\
& \leq \sum_{j=0}^{m-1} \alpha^j \beta^{n-j} B(n, j) (m-j) \frac{d_\infty(X, F(X))}{\rho} \\
& \quad + \sum_{j=m}^n \alpha^m \beta^{j-m} B(j-1, m-1) (n-j) \frac{d_\infty(X, F(X))}{\rho} \\
& \leq \left[ \lambda \sqrt{\left(n \frac{\tau}{\lambda} - m\right)^2 + n \frac{\tau}{\lambda} \frac{\lambda - \tau}{\lambda}} \right. \\
& \quad \left. + \tau \sqrt{\frac{\lambda^2}{\tau^2} \frac{\lambda - \tau}{\lambda} m + \left(\frac{\lambda}{\tau} \frac{\lambda - \tau}{\lambda} m + m - n\right)^2} \right] \frac{d_\infty(X, F(X))}{\rho} \\
& = \left[ \sqrt{(n\tau - m\lambda)^2 + n\tau(\lambda - \tau)} \right. \\
& \quad \left. + \sqrt{m\lambda(\lambda - \tau) + (m\lambda - n\tau)^2} \right] \frac{d_\infty(X, F(X))}{\rho}.
\end{aligned} \tag{41}$$

For  $\tau = \frac{t}{n}, \lambda = \frac{t}{m}$ , the above reads

$$d_\infty((J_{t/n,\rho})^n(X), (J_{t/m,\rho})^m(X)) \leq 2t \left| \frac{1}{n} - \frac{1}{m} \right|^{1/2} \frac{d_\infty(X, F(X))}{\rho},$$

so the limit in (36) exists by completeness and satisfies (37), moreover the above also yield the Lipschitz estimate (38). The rest of the properties is routine to prove, by following the steps of the proof of Theorem 4.6.  $\square$

**Lemma 6.8.** *Let  $F : \mathbb{P} \mapsto \mathbb{P}$  be a nonexpansive map,  $\rho > 0$  and let  $S_\rho(t)$  be the semigroup constructed in Theorem 6.7. Then for  $t > 0$ ,  $X \in \mathbb{P}$  and  $m \in \mathbb{N}$  we have*

$$\begin{aligned}
& \text{(i)} \quad S_\rho(t)X = S_1(t/\rho)X, \\
& \text{(ii)} \quad d_\infty(F^m(X), S_\rho(t)X) \leq \left[ \frac{t}{\rho} - m + 2\sqrt{\left(\frac{t}{\rho} - m\right)^2 + \frac{t}{\rho}} \right] d_\infty(X, F(X)).
\end{aligned}$$

*Proof.* The proof of (i) follows directly from the fact that for  $\rho, \lambda > 0$ , we have  $J_{\lambda,\rho} = J_{\lambda/\rho,1}$ .

We turn to the proof of (ii) which is more involved. Proposition 2.10 and (32) yields that

$$\begin{aligned}
d_\infty(F^m(X), (J_{t/n,1})^n(X)) & \leq \frac{1}{1 + \frac{t}{n}} d_\infty(F^m(X), (J_{t/n,1})^{n-1}(X)) \\
& \quad + \frac{\frac{t}{n}}{1 + \frac{t}{n}} d_\infty(F^m(X), F((J_{t/n,1})^n(X))).
\end{aligned}$$

Using the above inequality recursively, we get

$$\begin{aligned}
d_\infty(F^m(X), (J_{t/n,1})^n(X)) &\leq \left(1 + \frac{t}{n}\right)^{-n} d_\infty(F^m(X), X) \\
&\quad + \frac{t}{n} \sum_{k=1}^n \left(1 + \frac{t}{n}\right)^{-(n-k)} d_\infty(F^m(X), F((J_{t/n,1})^k(X))) \\
&\leq \left(1 + \frac{t}{n}\right)^{-n} m d_\infty(F(X), X) \\
&\quad + \frac{t}{n} \sum_{k=1}^n \left(1 + \frac{t}{n}\right)^{-(n-k)} d_\infty(F^{m-1}(X), (J_{t/n,1})^k(X)).
\end{aligned}$$

For  $n \in \mathbb{N}$  define

$$\begin{aligned}
f_n(s) &:= \sum_{k=1}^n \left(1 + \frac{t}{n}\right)^{-(n-k)} 1_{\left(\frac{(k-1)t}{n}, \frac{kt}{n}\right]}(s), \\
g_n(s) &:= \sum_{k=1}^n d_\infty(F^{m-1}(X), (J_{t/n,1})^k(X)) 1_{\left(\frac{(k-1)t}{n}, \frac{kt}{n}\right]}(s),
\end{aligned}$$

so that the above becomes

$$d_\infty(F^m(X), (J_{t/n,1})^n(X)) \leq \left(1 + \frac{t}{n}\right)^{-n} m d_\infty(F(X), X) + \int_0^t f_n(s) g_n(s) ds.$$

We will show that  $f_n(s) \rightarrow e^{-(t-s)}$  and  $g_n(s) \rightarrow d_\infty(F^{m-1}(X), S_1(s)X)$  for  $s \in (0, t]$  and that  $\sup_{n \in \mathbb{N}, s \in [0, t]} |f_n(s)| |g_n(s)| < \infty$ . Then by dominated convergence we will have

$$\begin{aligned}
d_\infty(F^m(X), S_1(t)X) &\leq e^{-t} m d_\infty(X, F(X)) \\
(42) \quad &\quad + \int_0^t e^{-(t-s)} d_\infty(F^{m-1}(X), S_1(s)X) ds.
\end{aligned}$$

Firstly it is routine to see that  $f_n(s) \rightarrow e^{-(t-s)}$ . To prove the other claim, let  $n \in \mathbb{N}$  and  $s \in (0, t]$ . There is a unique  $0 < k_{s,n} \leq n$  such that

$$\frac{(k_{s,n} - 1)t}{n} < s \leq \frac{k_{s,n}t}{n}.$$

Substituting  $k_{s,n}, \frac{t}{n}, m, \frac{s}{m}$  for  $n, \tau, m, \lambda$  respectively in (41) gives

$$\begin{aligned}
d_\infty((J_{t/n,1})^{k_{s,n}}(X), (J_{s/m,1})^m(X)) &\leq \left[ \sqrt{\left(\frac{k_{s,n}t}{n} - s\right)^2 + \frac{k_{s,n}t}{n} \left(\frac{s}{m} - \frac{t}{n}\right)} \right. \\
&\quad \left. + \sqrt{s \left(\frac{s}{m} - \frac{t}{n}\right) + \left(\frac{k_{s,n}t}{n} - s\right)^2} \right] 2d_\infty(X, F(X))
\end{aligned}$$

and taking the limit  $m \rightarrow \infty$  we get

$$(43) \quad \begin{aligned} d_\infty((J_{t/n,1})^{k_{s,n}}(X), S_1(s)(X)) &\leq \left[ \sqrt{\left(\frac{k_{s,n}t}{n} - s\right)^2 - \frac{k_{s,n}t}{n} \frac{t}{n}} \right. \\ &\quad \left. + \sqrt{\left(\frac{k_{s,n}t}{n} - s\right)^2 - s \frac{t}{n}} \right] 2d_\infty(X, F(X)). \end{aligned}$$

We also have

$$\begin{aligned} |g_n(s) - d_\infty(F^{m-1}(X), S_1(s)X)| &= |d_\infty(F^{m-1}(X), (J_{t/n,1})^{k_{s,n}}(X)) \\ &\quad - d_\infty(F^{m-1}(X), S_1(s)X)| \end{aligned}$$

which combined with (43) and the triangle inequality yields

$$g_n(s) \rightarrow d_\infty(F^{m-1}(X), S_1(s)X),$$

so (42) follows. Now if  $d_\infty(F(X), X) = 0$ , then  $J_{\lambda,\rho}(X) = X$  for all  $\lambda, \rho > 0$ , and we have  $S_\rho(s)X = X$  and (ii) follows. Assume  $d_\infty(F(X), X) > 0$  and for  $m \geq 0, s \geq 0$  let

$$\phi_m(s) := \frac{d_\infty(F^m(X), S_1(s)X)}{d_\infty(F(X), X)},$$

so that (42) gives

$$\begin{aligned} \phi_m(t) &\leq e^{-t}m + \int_0^t e^{-(t-s)}\phi_{m-1}(s)ds \\ &= e^{-t} \left( m + \int_0^t e^s \phi_{m-1}(s)ds \right), \end{aligned}$$

so that if  $f_m(s) := e^s \phi_m(s)$ , then

$$(44) \quad f_m(t) \leq m + \int_0^t f_{m-1}(s)ds.$$

It is straightforward to check that

$$f_m(t) = e^t(t - m) + 2 \sum_{j=0}^m (m - j) \frac{t^j}{j!}$$

satisfies the recursion (44) with equality, so that

$$\phi_m(t) \leq (t - m) + 2 \sum_{j=0}^m (m - j) \frac{t^j}{j!} e^{-t}.$$

In view of the estimate

$$\begin{aligned} \sum_{j=0}^{\infty} \frac{|j - m| m^j \alpha^j}{j!} &\leq e^{\frac{m\alpha}{2}} \sum_{j=0}^{\infty} \frac{(j - m)^2 m^j \alpha^j}{j!} \\ &= e^{m\alpha} [m^2(\alpha - 1)^2 + m(\alpha - 1) + m]^{1/2} \end{aligned}$$

from [22], we get that

$$\begin{aligned}
\phi_m(t) &\leq (t-m) + 2 \sum_{j=0}^m (j-m) \frac{t^j}{j!} e^{-t} \\
&\leq (t-m) + 2e^{-t} \sum_{j=0}^{\infty} \frac{|j-m| m^j \left(\frac{t}{m}\right)^j}{j!} \\
&\leq (t-m) + 2e^{-t} e^{\frac{t}{2}} \sum_{j=0}^{\infty} \frac{(j-m)^2 m^j \left(\frac{t}{m}\right)^j}{j!} \\
&= (t-m) + 2[(t-m)^2 + t]^{1/2}.
\end{aligned}$$

Thus, (ii) follows from the above combined with (i).  $\square$

**Proposition 6.9.** *For  $\rho > 0$  let  $F_\rho : \mathbb{P} \mapsto \mathbb{P}$  be a nonexpansive map, and let  $S_\rho(t)$  denote the semigroup generated by  $J_{\lambda,\rho}$  corresponding to the nonexpansive map  $F := F_\rho$  for each  $\rho > 0$  in Theorem 6.7. If*

$$J_{\lambda,\rho}(X) \rightarrow J_\lambda^\mu(X)$$

*in  $d_\infty$  as  $\rho \rightarrow 0+$  for a fixed  $\mu \in \mathcal{P}^1(\mathbb{P})$  and all  $X \in \mathbb{P}$ , then*

$$(45) \quad S_\rho(t)X \rightarrow S(t)X$$

*in  $d_\infty$  for all  $X \in \mathbb{P}$  as  $\rho \rightarrow 0+$ , where  $S(t)$  is the semigroup generated by  $J_\lambda^\mu$  in Theorem 4.6. Moreover the limit in (45) is uniform on compact time intervals.*

*Proof.* Fix a  $T > 0$ ,  $X \in \mathbb{P}$  and let  $0 < t < T$ . For all  $\lambda > 0$  the assumption implies

$$(46) \quad \frac{d_\infty(X, J_{\lambda,\rho}(X))}{\lambda} \rightarrow \frac{d_\infty(X, J_\lambda^\mu(X))}{\lambda} \leq \int_{\mathbb{P}} d_\infty(X, A) d\mu(A)$$

as  $\rho \rightarrow 0+$ , where the inequality follows from (23). We also have the following estimates

$$(47) \quad \begin{aligned} d_\infty(S_\rho(t)X, S(t)X) &\leq d_\infty(S_\rho(t)X, S_\rho(t)J_{\lambda,\rho}(X)) + d_\infty(S_\rho(t)J_{\lambda,\rho}(X), S(t)X) \\ &\leq d_\infty(X, J_{\lambda,\rho}(X)) + d_\infty(S_\rho(t)J_{\lambda,\rho}(X), S(t)X) \end{aligned}$$

and

$$\begin{aligned}
d_\infty(S_\rho(t)J_{\lambda,\rho}(X), S(t)X) &\leq d_\infty(S_\rho(t)J_{\lambda,\rho}(X), (J_{t/n,\rho})^n J_{\lambda,\rho}(X)) \\
&\quad + d_\infty((J_{t/n,\rho})^n J_{\lambda,\rho}(X), (J_{t/n,\rho})^n(X)) \\
(48) \quad &\quad + d_\infty((J_{t/n,\rho})^n(X), (J_{t/n}^\mu)^n(X)) \\
&\quad + d_\infty((J_{t/n}^\mu)^n(X), S(t)X).
\end{aligned}$$

We need to find upper bound on the terms in (48). Firstly by (37) we have

$$\begin{aligned}
d_\infty(S_\rho(t)J_{\lambda,\rho}(X), (J_{t/n,\rho})^n J_{\lambda,\rho}(X)) &\leq \frac{2t}{\sqrt{n}} \frac{d_\infty(J_{\lambda,\rho}(X), F_\rho(J_{\lambda,\rho}(X)))}{\rho} \\
&= \frac{2t}{\sqrt{n}} \frac{d_\infty(X, J_{\lambda,\rho}(X))}{\lambda}
\end{aligned}$$

where equality follows from (32). Since  $J_{\lambda,\rho}$  is a contraction, for all  $n \in \mathbb{N}$  we get

$$d_\infty((J_{t/n,\rho})^n J_{\lambda,\rho}(X), (J_{t/n,\rho})^n(X)) \leq d_\infty(J_{\lambda,\rho}(X), X)$$

and by (28) we get

$$d_\infty \left( \left( J_{t/n}^\mu \right)^n (X), S(t)X \right) \leq \frac{2t}{\sqrt{n}} \int_{\mathbb{P}} d_\infty(X, A) d\mu(A).$$

From the above we obtained the following:

$$(49) \quad \begin{aligned} d_\infty(S_\rho(t)X, S(t)X) &\leq 2d_\infty(J_{\lambda,\rho}(X), X) + \frac{2t}{\sqrt{n}} \frac{d_\infty(X, J_{\lambda,\rho}(X))}{\lambda} \\ &\quad + \frac{2t}{\sqrt{n}} \int_{\mathbb{P}} d_\infty(X, A) d\mu(A) \\ &\quad + d_\infty((J_{t/n,\rho})^n(X), (J_{t/n}^\mu)^n(X)). \end{aligned}$$

Now let  $\epsilon > 0$ . Choose  $\lambda_0 > 0$  so that  $\lambda_0 \int_{\mathbb{P}} d_\infty(X, A) d\mu(A) < \epsilon$ . By (46), there exists a  $\delta > 0$  such that for  $\rho < \delta$  we have

$$\frac{d_\infty(X, J_{\lambda_0,\rho}(X))}{\lambda_0} \leq \frac{d_\infty(X, J_{\lambda_0}^\mu(X))}{\lambda_0} + \epsilon.$$

Thus,

$$d_\infty(X, J_{\lambda_0,\rho}(X)) \leq \lambda_0 \int_{\mathbb{P}} d_\infty(X, A) d\mu(A) + \epsilon \lambda_0 < 2\epsilon$$

for  $\rho < \delta$ . Next, choose an  $n_0 \in \mathbb{N}$  such that

$$\frac{2t}{\sqrt{n}} \left[ \frac{d_\infty(X, J_{\lambda_0,\rho}(X))}{\lambda} + \int_{\mathbb{P}} d_\infty(X, A) d\mu(A) \right] < \epsilon$$

for all  $n \geq n_0$  and  $t < T$ .

Finally for  $t < T$  we estimate

$$\begin{aligned} d_\infty((J_{t/n_0,\rho})^{n_0}(X), (J_{t/n_0}^\mu)^{n_0}(X)) &\leq d_\infty((J_{t/n_0,\rho})^{n_0}(X), (J_{t/n_0,\rho})^{n_0-1} J_{t/n_0}^\mu(X)) \\ &\quad + d_\infty((J_{t/n_0,\rho})^{n_0-1} J_{t/n_0}^\mu(X), (J_{t/n_0}^\mu)^{n_0-1} J_{t/n_0}^\mu(X)) \end{aligned}$$

which, by induction together with the assumption of the assertion, yields the existence of  $\rho_0 > 0$  such that for  $\rho < \rho_0$  we get that

$$d_\infty((J_{t/n_0,\rho})^{n_0}(X), (J_{t/n_0}^\mu)^{n_0}(X)) < \epsilon.$$

Combining the above with (49) we obtain (45).

To show that for a fixed  $X \in \mathbb{P}$  the convergence in (45) is uniform for  $t < T$ , pick  $\tau \in (0, T)$ . By the triangle inequality and contraction property of  $S_\rho(t), S(t)$ , (38) and (29) we get

$$\begin{aligned} d_\infty(S(t)X, S_\rho(t)X) &\leq d_\infty(S(t)X, S_\rho(t)J_{\lambda,\rho}(X)) + d_\infty(S_\rho(t)J_{\lambda,\rho}(X), S_\rho(t)X) \\ &\leq d_\infty(S(t)X, S(\tau)X) + d_\infty(S(\tau)X, S_\rho(\tau)J_{\lambda,\rho}(X)) \\ &\quad + d_\infty(S_\rho(\tau)J_{\lambda,\rho}(X), S_\rho(t)J_{\lambda,\rho}(X)) + d_\infty(J_{\lambda,\rho}(X), X) \\ &\leq 2|t - \tau| \int_{\mathbb{P}} d_\infty(X, A) d\mu(A) + d_\infty(S(\tau)X, S_\rho(\tau)J_{\lambda,\rho}(X)) \\ &\quad + 2|t - \tau| \frac{d_\infty(J_{\lambda,\rho}(X), F_\rho(J_{\lambda,\rho}(X)))}{\rho} + d_\infty(J_{\lambda,\rho}(X), X). \end{aligned}$$

Now as in the first part, we can fix a  $\lambda > 0$  such that for sufficiently small  $\rho > 0$  the quantity  $d_\infty(J_{\lambda,\rho}(X), X)$  becomes arbitrarily small. We again have that  $\frac{d_\infty(J_{\lambda,\rho}(X), F_\rho(J_{\lambda,\rho}(X)))}{\rho} = \frac{d_\infty(J_{\lambda,\rho}(X), X)}{\lambda}$ , and for fixed  $\lambda > 0$ , this term is bounded as

$\rho \rightarrow 0+$ . We have also seen before that  $d_\infty(S(\tau)X, S_\rho(\tau)J_{\lambda,\rho}(X))$  is small for small  $\rho > 0$ . Now we can use the compactness of  $[0, T]$  to conclude that the convergence is uniform on  $[0, T]$  in (45).  $\square$

**Theorem 6.10.** *For each  $\rho > 0$  let  $F_\rho : \mathbb{P} \mapsto \mathbb{P}$  be a nonexpansive map and let  $J_{\lambda,\rho}$  be the resolvent generated by  $F_\rho$  in (32) for each  $\rho > 0$ . If*

$$J_{\lambda,\rho}(X) \rightarrow J_\lambda^\mu(X)$$

*in  $d_\infty$  as  $\rho \rightarrow 0+$  for a fixed  $\mu \in \mathcal{P}^1(\mathbb{P})$  and all  $X \in \mathbb{P}$ , then*

$$(50) \quad (F_{\frac{t}{n}})^n(X) \rightarrow S(t)X$$

*in  $d_\infty$  for all  $X \in \mathbb{P}$  as  $n \rightarrow \infty$ , where  $S(t)$  is the semigroup generated by  $J_\lambda^\mu$  in Theorem 4.6. Moreover the limit in (50) is uniform on compact time intervals.*

*Proof.* Fix  $T > 0$ , let  $X \in \mathbb{P}$  and let  $0 < t \leq T$ . Let  $S_\rho(t)$  denote the semigroup generated by  $J_{\lambda,\rho}$ . We have

$$(51) \quad d_\infty(S(t)X, (F_{\frac{t}{n}})^n(X)) \leq d_\infty(S(t)X, S_{\frac{t}{n}}(t)X) + d_\infty(S_{\frac{t}{n}}(t)X, (F_{\frac{t}{n}})^n(X)).$$

For  $\rho > 0$ ,  $n \in \mathbb{N}$  and  $\lambda > 0$  we have

$$(52) \quad \begin{aligned} d_\infty(S_\rho(n\rho)X, (F_\rho)^n(X)) &\leq d_\infty(S_\rho(n\rho)X, S_\rho(n\rho)J_{\lambda,\rho}(X)) \\ &\quad + d_\infty(S_\rho(n\rho)J_{\lambda,\rho}(X), (F_\rho)^n(J_{\lambda,\rho}(X))) \\ &\quad + d_\infty((F_\rho)^n(J_{\lambda,\rho}(X)), (F_\rho)^n(X)) \\ &\leq 2d_\infty(J_{\lambda,\rho}(X), X) \\ &\quad + d_\infty(S_\rho(n\rho)J_{\lambda,\rho}(X), (F_\rho)^n(J_{\lambda,\rho}(X))). \end{aligned}$$

For  $\rho = \frac{t}{n}$  and  $\lambda > 0$  we have by Lemma 6.8 that

$$(53) \quad \begin{aligned} &d_\infty(S_\rho(n\rho)J_{\lambda,\rho}(X), (F_\rho)^n(J_{\lambda,\rho}(X))) \\ &\leq 2\sqrt{\left(n - \frac{n\rho}{\rho}\right)^2 + \frac{n\rho}{\rho}d_\infty(J_{\lambda,\rho}(X), F_\rho(J_{\lambda,\rho}(X)))} \\ &= 2\sqrt{n\rho} \frac{d_\infty(J_{\lambda,\rho}(X), X)}{\lambda} \leq 2\frac{T}{\sqrt{n}} \frac{d_\infty(J_{\lambda,\rho}(X), X)}{\lambda} \end{aligned}$$

where the equality follows from (32).

We have already seen in the proof of Proposition 6.9 after (49) how the convergence of the resolvents  $J_{\lambda,\rho}(X) \rightarrow J_\lambda^\mu(X)$  imply estimates on  $d_\infty(J_\lambda^\mu(X), X)$ ,  $d_\infty(J_{\lambda,\rho}(X), X)$  and  $\frac{d_\infty(J_{\lambda,\rho}(X), X)}{\lambda}$ . Thus the estimates (51), (52) and (53) along with Proposition 6.9 implies uniform convergence in (50) on compact time intervals.  $\square$

## 7. CONVERGENCE OF RESOLVENTS

In this section we prove the convergence of the resolvents  $J_{\lambda,\rho}(X) \rightarrow J_\lambda^\mu(X)$  in  $d_\infty$  for finitely supported measures  $\mu = \sum_{i=1}^n \frac{1}{n} \delta_{A_i} \in \mathcal{P}^1(\mathbb{P})$  in Theorem 7.8. This result will play a key role later in proving a continuous time law of large numbers type result for  $\Lambda$ . The analysis of this section also proves the norm convergence of power means to the Karcher mean solving this conjecture mentioned in [18].

Let  $\mathbb{P}^*$  denote the dual cone of  $\mathbb{P}$ , i.e. the cone of all non-negative norm continuous linear functionals on  $\mathbb{P}$ .

**Lemma 7.1.** *Let  $A, B \in \mathbb{P}$ . Then there exists an  $\omega \in \mathbb{P}^*$  with  $\omega(I) = 1$  such that either*

$$e^{d_\infty(B \#_t A, B)} \omega(B) = \omega(B \#_t A)$$

*or*

$$\omega(B \#_t A) = e^{d_\infty(B \#_t A, B)} \omega(B)$$

*holds for all  $t \in [0, 1]$ .*

*Proof.* By definition we have that

$$d_\infty(A, B) = \log \max\{\inf\{\alpha > 0 : A \leq \alpha B\}, \inf\{\alpha > 0 : B \leq \alpha A\}\}.$$

Assume first that  $B = I$ . Then

$$e^{d_\infty(I \#_t A, I)} = \max\{\|A^t\|, \|A^{-t}\|\},$$

since we have

$$\inf\{\alpha > 0 : A \leq \alpha I\} = \|A\|.$$

We also have that there exists a net  $v_\alpha \in \mathcal{H}$  with  $|v_\alpha| = 1$  and  $\lim_{\alpha \rightarrow \infty} |Av_\alpha| = \|A\|$ . That is

$$\|A\|^2 = \lim_{\alpha \rightarrow \infty} v_\alpha^* A^* A v_\alpha = \lim_{\alpha \rightarrow \infty} v_\alpha^* A^2 v_\alpha = \lim_{\alpha \rightarrow \infty} \omega_\alpha(A^2),$$

where  $v_\alpha^*(\cdot)v_\alpha =: \omega_\alpha \in \mathbb{P}^*$  and  $\omega_\alpha(I) = \|\omega_\alpha\|_* = 1$ . Since the convex set  $\{\nu \in \mathbb{P}^* : \nu(I) = \|\nu\|_* = 1\}$  is weak-\* compact, there exists a subnet of  $\omega_\alpha$  again denoted by  $\omega_\alpha$  that has a limit point  $\omega \in \{\nu \in \mathbb{P}^* : \nu(I) = \|\nu\|_* = 1\}$ . Then the state  $\omega$  satisfies  $\|A\|^2 = \omega(A^2)$  which is equivalent to  $\|A\| = \omega(A)$ , more generally by the monotonicity of the power function it follows that

$$\|A^t\| = \omega(A^t).$$

If  $e^{d_\infty(A, I)} = \|A\|$ , this yields that

$$\omega(I \#_t A) = \omega(A^t) = e^{d_\infty(I \#_t A, I)} \omega(I).$$

In the other case when  $e^{d_\infty(A, I)} = \|A^{-1}\|$  by the same argument as above, we can find an  $\omega \in \{\nu \in \mathbb{P}^* : \nu(I) = \|\nu\|_* = 1\}$  such that

$$e^{-d_\infty(I \#_t A, I)} \omega(I) = \omega(I \#_t A).$$

Now the case of arbitrary  $B \in \mathbb{P}$  follows by considering first

$$\omega\left(\left(B^{-1/2}AB^{-1/2}\right)^t\right) = e^{d_\infty\left(\left(B^{-1/2}AB^{-1/2}\right)^t, I\right)} \omega(I)$$

which is equivalent to

$$\hat{\omega}(B \#_t A) = e^{d_\infty(I \#_t (B^{-1/2}AB^{-1/2}), I)} \hat{\omega}(B) = e^{d_\infty(B \#_t A, B)} \hat{\omega}(B)$$

where  $\hat{\omega}(X) := \frac{1}{\omega(B^{-1})} \omega(B^{-1/2}XB^{-1/2})$ . The other equality in the assertion follows similarly from the case  $B = I$ .  $\square$

We will use the notation

$$\phi_\mu(X) := \int_{\mathbb{P}} \log_X A d\mu(A)$$

for  $\mu \in \mathcal{P}^1(\mathbb{P})$  and  $X \in \mathbb{P}$ .

**Proposition 7.2.** *Let  $\mu \in \mathcal{P}^1(\mathbb{P})$  and  $X \in \mathbb{P}$ . Let  $D\phi_\mu(X)[V]$  denote the Fréchet-derivative of  $\phi_\mu$  in the direction  $V \in \mathbb{S}$ . Then the linear map  $D\phi_\mu(\Lambda(\mu)) : \mathbb{S} \mapsto \mathbb{S}$  is injective, in particular*

$$(54) \quad \frac{1}{\|\Lambda(\mu)^{-1}\| \|\Lambda(\mu)\|} \leq \|D\phi_\mu(\Lambda(\mu))\|$$

where  $\|D\phi_\mu(\Lambda(\mu))\| := \sup_{V \in \mathbb{S}, \|V\|=1} \|D\phi_\mu(\Lambda(\mu))[V]\|$ .

*Proof.* Let  $X, Y \in \mathbb{P}$  and according to (25) let  $\gamma(t) := S(t)X$ ,  $\eta(t) := S(t)Y$ . Then by (26) we have

$$(55) \quad d_\infty(\gamma(t), \eta(t)) \leq e^{-t} d_\infty(X, Y).$$

Then by Lemma 7.1 there exists an  $\omega \in \mathbb{P}^*$  with  $\omega(I) = 1$  such that either

$$(56) \quad \omega(X) = e^{d_\infty(X, Y)} \omega(Y)$$

or

$$(57) \quad \omega(Y) = e^{d_\infty(X, Y)} \omega(X).$$

Assume that (56) holds. In general we have that

$$\gamma(t) \leq e^{d_\infty(\gamma(t), \eta(t))} \eta(t)$$

which combined with (55) yields

$$\gamma(t) \leq e^{e^{-t} d_\infty(X, Y)} \eta(t).$$

Thus, since  $\omega$  is positive we have

$$(58) \quad \omega(\gamma(t)) \leq e^{e^{-t} d_\infty(X, Y)} \omega(\eta(t))$$

where for  $t = 0$  we have equality by (56). Hence we may take the derivative of (58) at  $t = 0$  to get

$$(59) \quad \omega(\dot{\gamma}(0)) \leq e^{d_\infty(X, Y)} \omega(\dot{\eta}(0)) - d_\infty(X, Y) e^{d_\infty(X, Y)} \omega(\eta(0)).$$

If (57) holds then we start from

$$\eta(t) \leq e^{d_\infty(\gamma(t), \eta(t))} \gamma(t)$$

to obtain

$$(60) \quad \omega(\dot{\eta}(0)) \leq e^{d_\infty(X, Y)} \omega(\dot{\gamma}(0)) - d_\infty(X, Y) e^{d_\infty(X, Y)} \omega(\gamma(0)).$$

by a similar argument.

Now assume that  $\eta(0) = \Lambda(\mu)$  and  $\gamma(0) = \eta(0) \#_s Z$  for  $Z \in \mathbb{P}$  and  $s \in [0, 1]$ . Then by Theorem 4.8 we have that

$$(61) \quad \begin{aligned} \dot{\eta}(0) &= \phi_\mu(\Lambda(\mu)) = 0, \\ \dot{\gamma}(0) &= \phi_\mu(\gamma(0)). \end{aligned}$$

First assume that (59) holds so that by (61) and Lemma 7.1 we have

$$\begin{aligned} \omega(\phi_\mu(\gamma(0))) &\leq -d_\infty(\eta(0), \eta(0) \#_s Z) e^{d_\infty(\eta(0), \eta(0) \#_s Z)} \omega(\eta(0)) \\ &= -s d_\infty(\Lambda(\mu), Z) e^{s d_\infty(\Lambda(\mu), Z)} \omega(\Lambda(\mu)) \end{aligned}$$

for all  $s \in [0, 1]$ . Since  $\phi_\mu(\gamma(0)) = \phi_\mu(\Lambda(\mu) \#_s Z)$ , the above yields

$$(62) \quad \omega(\phi_\mu(\Lambda(\mu) \#_s Z)) \leq -s d_\infty(\Lambda(\mu), Z) e^{s d_\infty(\Lambda(\mu), Z)} \omega(\Lambda(\mu)).$$



For  $s = 0$  we have equality in (62) since the two sides both equal to 0, thus we can differentiate (62) at  $s = 0$  to get

$$\omega(D\phi_\mu(\Lambda(\mu))[\log_{\Lambda(\mu)}(Z)]) \leq -d_\infty(\Lambda(\mu), Z)\omega(\Lambda(\mu))$$

and it follows that

$$(63) \quad \omega(\Lambda(\mu)) \leq -\omega\left(D\phi_\mu(\Lambda(\mu))\left[\frac{1}{\|\log(\Lambda(\mu)^{-1/2}Z\Lambda(\mu)^{-1/2})\|}\log_{\Lambda(\mu)}(Z)\right]\right).$$

In the other case when (60) holds, by a similar argument we obtain

$$d_\infty(\Lambda(\mu), Z)\omega(\Lambda(\mu)) \leq \omega(D\phi_\mu(\Lambda(\mu))[\log_{\Lambda(\mu)}(Z)])$$

and thus

$$(64) \quad \omega(\Lambda(\mu)) \leq \omega\left(D\phi_\mu(\Lambda(\mu))\left[\frac{1}{\|\log(\Lambda(\mu)^{-1/2}Z\Lambda(\mu)^{-1/2})\|}\log_{\Lambda(\mu)}(Z)\right]\right).$$

As  $Z$  ranges over  $\mathbb{P}$ , the expression  $\log_{\Lambda(\mu)}(Z)$  attains all possible values in  $\mathbb{S}$  for which either we have (63) or (64) moreover  $\omega(\Lambda(\mu)) > 0$  since  $\omega \in \mathbb{P}^*$  and  $\omega(I) = 1$  and there exists  $\epsilon > 0$  such that  $\epsilon I \leq \Lambda(\mu)$ , thus  $D\phi_\mu(\Lambda(\mu)) : \mathbb{S} \mapsto \mathbb{S}$  is an injective bounded linear map. The inequality (54) follows (63) or (64).  $\square$

We need a few facts from the theory of bounded linear operators. We denote the predual of the von Neumann algebra  $\mathcal{B}(\mathcal{H})$  by  $\mathcal{B}(\mathcal{H})_*$  which is the ideal of trace class operators on  $\mathcal{H}$ . If we restrict to the self-adjoint part  $\mathbb{S}$ , then  $\mathbb{S}$  is a real Banach space with predual  $\mathbb{S}_* := \{X \in \mathcal{B}(\mathcal{H})_*, X^* = X\}$  and dual space  $\mathbb{S}^* := \{X \in \mathcal{B}(\mathcal{H})^*, X^* = X\}$  where  $\mathcal{B}(\mathcal{H})^*$  denotes the predual of the *universal enveloping von Neumann algebra*  $\mathcal{B}(\mathcal{H})^{**}$ . Since  $\mathbb{S}^*$  is the self-adjoint part of the ideal of trace-class operators  $\mathcal{B}(\mathcal{H})^*$  in the universal enveloping von Neumann algebra  $\mathcal{B}(\mathcal{H})^{**}$  and thus is the unique predual of the von Neumann algebra  $\mathcal{B}(\mathcal{H})^{**}$ , we have that any given  $X \in \mathbb{S}^*$  can be uniquely decomposed as  $X = X^+ - X^-$  where  $X^+, X^- \geq 0$  and such that the support projections of  $X^+$  and  $X^-$  are orthogonal by Theorem III.4.2. in [30]. The locally convex topology  $\sigma(\mathbb{S}_*, \mathbb{S})$  is called the  $\sigma$ -weak or ultraweak operator topology on  $\mathbb{S}$ .

**Lemma 7.3.** *Let  $\mu \in \mathcal{P}^1(\mathbb{P})$  and  $X \in \mathbb{P}$ . Then the linear map  $D\phi_\mu(\Lambda(\mu)) : \mathbb{S} \mapsto \mathbb{S}$  has dense range in the ultraweak operator topology.*

*Proof.* Firstly, notice that for  $X, A \in \mathbb{P}$  and invertible  $C \in \mathcal{B}(\mathcal{H})$  we have

$$\begin{aligned} C \log_X AC^* &= CX \log(X^{-1}A)C^* \\ &= CXC^*C^{-*} \log(X^{-1}A)C^* \\ &= CXC^* \log(C^{-*}X^{-1}C^{-1}CAC^*) \\ &= \log_{CXC^*}(CAC^*), \end{aligned}$$

hence  $CD \log_X A[V]C^* = D \log_{CXC^*}(CAC^*)[CVC^*]$ , where the differentiation is with respect to the variable  $X$  and  $A$  is fixed. Thus for arbitrary  $V \in \mathbb{S}$  it also follows that

$$D\phi_\mu(\Lambda(\mu))[V] = \Lambda(\mu)^{1/2}D\phi_{\hat{\mu}}(I)[\Lambda(\mu)^{-1/2}V\Lambda(\mu)^{-1/2}]\Lambda(\mu)^{1/2}$$

where  $d\hat{\mu}(A) := d\mu(\Lambda(\mu)^{1/2}A\Lambda(\mu)^{1/2})$  and  $\Lambda(\hat{\mu}) = I$  as well. Thus it is enough to prove that the range of  $D\phi_\mu(\Lambda(\mu))$  is dense when  $\Lambda(\mu) = I$ .

So without loss of generality assume that  $\Lambda(\mu) = I$ . It is well known that on a locally convex space  $X$ , a linear operator  $T : X \mapsto X$  has dense range if and

only if there are no nonzero linear functionals in the dual space  $X^*$  which vanish on the range of the map  $T$ . So assume on the contrary that there exists a nonzero ultraweakly continuous linear functional  $\tau \in \mathbb{S}_*$  such that  $\tau(D\phi_\mu(\Lambda(\mu))[V]) = 0$  for all  $V \in \mathbb{S}$ . In what follows we will reverse the construction given in the proof of Lemma 7.1. Consider the unique decomposition  $\tau = \tau_+ - \tau_-$  where  $\tau_+, \tau_- \geq 0$  and the support projections  $s(\tau_+), s(\tau_-)$  of  $\tau_+, \tau_-$  are orthogonal. Let  $X := \exp(s(\tau_+) - s(\tau_-))$ . Then by the orthogonality of  $s(\tau_+), s(\tau_-)$  we have

$$X = \exp(s(\tau_+)) \oplus \exp(-s(\tau_-)) \oplus I_{E/(\text{Rg}(s(\tau_+)) \cup \text{Rg}(s(\tau_-)))}$$

if we restrict the domain of  $s(\tau_+), s(\tau_-)$  to their range respectively. Consider the states  $\hat{\tau}_+, \hat{\tau}_- \in \mathbb{S}^*$  defined as

$$\begin{aligned}\hat{\tau}_+(\cdot) &:= \frac{1}{\tau_+(I)} \tau_+(\cdot), \\ \hat{\tau}_-(\cdot) &:= \frac{1}{\tau_-(I)} \tau_-(\cdot).\end{aligned}$$

By construction it follows that they are both norming linear functionals for  $X$  in the following sense:

$$\begin{aligned}\hat{\tau}_+(X) &= e, \\ \hat{\tau}_-(X^{-1}) &= e.\end{aligned}$$

So we can follow the argumentation from (56) with  $Y = I$  and the state  $\omega := \hat{\tau}_+$  to arrive at (63), that is

$$(65) \quad \hat{\tau}_+(I) \leq -\hat{\tau}_+ \left( D\phi_\mu(I) \left[ \frac{1}{\|\log(X)\|} \log(X) \right] \right).$$

Similarly, in the other case we choose  $\omega := \hat{\tau}_-$  in (57) to obtain (64), which is

$$(66) \quad \hat{\tau}_-(I) \leq \hat{\tau}_- \left( D\phi_\mu(I) \left[ \frac{1}{\|\log(X)\|} \log(X) \right] \right).$$

We have that  $\tau = \tau_+ - \tau_-$ , hence (65) and (66) yields

$$0 < \tau_+(I)\hat{\tau}_+(I) + \tau_-(I)\hat{\tau}_-(I) \leq -\tau \left( D\phi_\mu(I) \left[ \frac{1}{\|\log(X)\|} \log(X) \right] \right)$$

contradicting the initial assumption  $\tau(D\phi_\mu(\Lambda(\mu))[V]) = 0$  for all  $V \in \mathbb{S}$ .  $\square$

**Lemma 7.4.** *Let  $\mu \in \mathcal{P}^1(\mathbb{P})$  and  $X \in \mathbb{P}$ . Let  $\pi : \mathcal{B}(\mathcal{H}) \mapsto \mathcal{A}$  be a unital  $*$ -representation into a unital  $C^*$ -algebra  $\mathcal{A}$ . Then the linear map  $D\phi_\mu(\Lambda(\mu)) : \mathbb{S} \mapsto \mathbb{S}$  commutes with  $\pi$ , i.e.*

$$\pi(D\phi_\mu(\Lambda(\mu))[V]) = D\phi_{\hat{\mu}}(\Lambda(\hat{\mu}))[\pi(V)]$$

where  $d\hat{\mu}(A) := d\mu(\pi^{-1}(A))$ . Moreover the linear map  $D\phi_\mu(\Lambda(\mu)) : \mathbb{S} \mapsto \mathbb{S}$  is ultraweakly continuous.

*Proof.* Firstly, since  $\pi$  is a  $*$ -representation it is automatically norm continuous, hence  $d\hat{\mu}(A)$  is well defined by the inverse image of the continuous map  $\pi$ . Secondly the continuous function  $\log$  can be defined as

$$\log(X) = \int_0^\infty \frac{\lambda}{\lambda^2 + 1} I - (\lambda I + X)^{-1} d\lambda,$$

thus

$$(67) \quad D \log(X)[V] = \int_0^\infty (\lambda I + X)^{-1} V (\lambda I + X)^{-1} d\lambda,$$

where both integrals converge in the norm topology. Then it is easy to see that

$$\pi(D \log(X)[V]) = D \log(\pi(X))[\pi(V)]$$

and similarly  $\Lambda(\cdot)$  and  $D\phi_\mu(\Lambda(\mu))[\cdot]$  commutes with  $\pi$  as well.

The ultraweak continuity of  $D\phi_\mu(\Lambda(\mu)) : \mathbb{S} \mapsto \mathbb{S}$  can be deduced from the formula

$$\begin{aligned} D\phi_\mu(\Lambda(\mu))[V] &= -\Lambda(\mu) \int_{\mathbb{P}} \int_0^\infty (\lambda I + \Lambda(\mu)^{-1} A)^{-1} \Lambda(\mu)^{-1} V \Lambda(\mu)^{-1} A \\ &\quad \times (\lambda I + \Lambda(\mu)^{-1} A)^{-1} d\lambda d\mu(A) \end{aligned}$$

which is derived using (67).  $\square$

**Proposition 7.5.** *Let  $\mu \in \mathcal{P}^1(\mathbb{P}(\mathcal{H}))$  and  $X \in \mathbb{P}(\mathcal{H})$ . Then the linear map  $D\phi_\mu(\Lambda(\mu)) : \mathbb{S}(\mathcal{H}) \mapsto \mathbb{S}(\mathcal{H})$  is a Banach space isomorphism.*

*Proof.* By Proposition 7.2 we know that  $D\phi_\mu(\Lambda(\mu))$  is injective and bounded below, hence its range is norm closed. As before, we assume without loss of generality that  $\Lambda(\mu) = I$ . Thus it remains to show that  $D\phi_\mu(I) : \mathbb{S}(\mathcal{H}) \mapsto \mathbb{S}(\mathcal{H})$  has norm dense range.

So, on the contrary assume that the range of  $D\phi_\mu(I) : \mathbb{S}(\mathcal{H}) \mapsto \mathbb{S}(\mathcal{H})$  is not norm dense. Then there exists a nonzero norm continuous linear functional  $\omega \in \mathbb{S}(\mathcal{H})^*$  such that

$$(68) \quad \omega(D\phi_\mu(I)[V]) = 0$$

for all  $V \in \mathbb{S}(\mathcal{H})$ . Let  $\pi_u : \mathcal{B}(\mathcal{H}) \mapsto \mathcal{B}(\mathcal{H}_{\pi_u})$  denote the *universal representation* on the GNS direct sum Hilbert space  $\mathcal{H}_{\pi_u}$ . Notice that  $\pi_u$  is unital. Notice that  $\omega \in \mathcal{B}(\mathcal{H}_{\pi_u})_*$ , hence by Lemma 7.4 we have

$$\begin{aligned} (69) \quad 0 &= \omega(D\phi_\mu(I)[V]) = \text{tr} \{ \omega \pi_u(D\phi_\mu(I)[V]) \} \\ &= \text{tr} \{ \omega D\phi_{\hat{\mu}}(\pi_u(I))[\pi_u(V)] \} \\ &= \text{tr} \{ \omega D\phi_{\hat{\mu}}(I)[\pi_u(V)] \} \end{aligned}$$

where  $d\hat{\mu}(A) := d\mu(\pi_u^{-1}(A))$ . We also know that the range of  $\pi_u$  is ultraweakly dense in the universal enveloping von Neumann algebra  $\mathcal{B}(\mathcal{H})^{**}$ , hence by the ultraweak continuity in Lemma 7.4 the map  $D\phi_{\hat{\mu}}(I)[\pi_u(\cdot)] : \mathbb{S}(\mathcal{H}) \mapsto \mathbb{S}(\mathcal{H}_{\pi_u})$  ultraweak continuously extends to the linear map  $D\phi_{\hat{\mu}}(I) : \mathbb{S}(\mathcal{H}_{\pi_u}) \mapsto \mathbb{S}(\mathcal{H}_{\pi_u})$ . Then by the ultraweak continuity of  $\omega$  on  $\mathcal{B}(\mathcal{H}_{\pi_u})$  and (69) we get that

$$\omega(D\phi_{\hat{\mu}}(I)[Z]) = \text{tr} \{ \omega D\phi_{\hat{\mu}}(I)[Z] \} = 0$$

for all  $Z \in \mathbb{S}(\mathcal{H}_{\pi_u})$ . This means that  $\omega$  on  $\mathbb{S}(\mathcal{H}_{\pi_u})$  is a nonzero ultraweakly continuous linear functional vanishing on the range of  $D\phi_{\hat{\mu}}(I) : \mathbb{S}(\mathcal{H}_{\pi_u}) \mapsto \mathbb{S}(\mathcal{H}_{\pi_u})$  contradicting the ultraweak density of the range of  $D\phi_{\hat{\mu}}(I)$  proved in Lemma 7.3.  $\square$

As a warm-up to more involved computations to follow we prove the norm convergence conjecture of the power means to the Karcher mean mentioned first in [18]. The conjecture states that  $\lim_{t \rightarrow 0+} P_t(\mu) = \Lambda(\mu)$  in the norm topology, where  $\mu \in \mathcal{P}^1(\mathbb{P})$  is such that the integral in (11) is bounded for all  $t \in [0, 1]$ . That is the case if all moments  $\int_{\mathbb{P}} d_{\infty}^p(X, A) d\mu(A) < +\infty$  for all  $p \geq 1$  and  $X \in \mathbb{P}$ .

**Lemma 7.6.** *Let  $\mu \in \mathcal{P}^1(\mathbb{P})$  with  $\int_{\mathbb{P}} d_{\infty}^p(X, A) d\mu(A) < +\infty$  for all  $p \geq 1$  and  $X \in \mathbb{P}$ . Consider the function  $F : [-1, 1] \times \mathbb{P} \mapsto \mathbb{S}$  defined as*

$$(70) \quad F(t, X) := \begin{cases} \int_{\mathbb{P}} \frac{1}{t} [X \#_t A - X] d\mu(A), & \text{if } t \neq 0, \\ \int_{\mathbb{P}} \log_X A d\mu(A), & \text{if } t = 0. \end{cases}$$

*Then  $F$  and its Fréchet derivative  $DF[\cdot]$  with respect to the variable  $X$  is a norm continuous function if we equip the product space  $[-1, 1] \times \mathbb{P}$  with the max norm generated by the individual Banach space norms on each factor.*

*Proof.* The function  $F$  is a smooth function everywhere except when  $t = 0$ , so we have to consider only this case.

The norm continuity of  $F$  follows easily from the fact that

$$\lim_{t \rightarrow 0} \frac{1}{t} [X \#_t A - X] = \log_X A.$$

We also have for  $t \neq 0$  and  $V \in \mathbb{S}$  that

$$\begin{aligned} D \left( \frac{1}{t} [X \#_t A - X] \right) [V] &= \frac{1}{t} D \left( X (X^{-1} A)^t - X \right) \\ &= \frac{1}{t} \left\{ V ((X^{-1} A)^t - I) \right. \\ &\quad \left. + X D \exp(t \log(X^{-1} A)) [t D \log(X^{-1} A) [-X^{-1} V X^{-1} A]] \right\} \\ &= V \frac{1}{t} ((X^{-1} A)^t - I) \\ &\quad + X D \exp(t \log(X^{-1} A)) [D \log(X^{-1} A) [-X^{-1} V X^{-1} A]]. \end{aligned}$$

Taking the limit  $t \rightarrow 0$  in the above, we obtain

$$\lim_{t \rightarrow 0} D \left( \frac{1}{t} [X \#_t A - X] \right) [V] = V \log(X^{-1} A) + X D \log(X^{-1} A) [-X^{-1} V X^{-1} A],$$

i.e. we derived that

$$D \log_X A [V] = \lim_{t \rightarrow 0} D \left( \frac{1}{t} [X \#_t A - X] \right) [V]$$

where differentiation is with respect to the variable  $X$ . Moreover it is easy to see, that the limit above is uniform for  $\|V\| \leq 1$ .  $\square$

**Theorem 7.7** (Continuity of  $P_t$ ). *Let  $\mu \in \mathcal{P}^1(\mathbb{P})$  with  $\int_{\mathbb{P}} d_{\infty}^p(X, A) d\mu(A) < +\infty$  for all  $p \geq 1$  and  $X \in \mathbb{P}$ . Then the family  $P_t(\mu)$  is norm continuous in  $t \in [-1, 1]$ , in particular*

$$(71) \quad \Lambda(\mu) = \lim_{t \rightarrow 0} P_t(\mu)$$

*in norm.*

*Proof.* We will use the Banach space version of the implicit function theorem, see for instance Theorem 4.9.3 in [21].

Consider the one-parameter family of functions  $F : [-1, 1] \times \mathbb{P} \mapsto \mathbb{S}$  defined in (70). Then by Lemma 7.6 the function  $F$  and its Fréchet derivative  $DF$  with respect to the second variable is continuous in the norm topology. Moreover  $F(0, \Lambda(\mu)) = 0$  and  $DF(0, \Lambda(\mu))[0, \cdot]$  is a Banach space isomorphism by Proposition 7.5. Therefore by the implicit function theorem (Theorem 4.9.3 [21]) there exists an open interval

$(a, -a)$  of  $0 \in [-1, 1]$  and a norm continuous function  $\hat{P}(t)$  such that the operator equation

$$(72) \quad F(t, \hat{P}(t)) = 0$$

is satisfied on  $(-a, a)$ , moreover it is uniquely satisfied there by the function  $\hat{P}(t)$ . Thus it follows by Proposition 2.6 that  $\hat{P}(t) = P_t(\mu)$  and also  $\hat{P}(0) = \Lambda(\mu)$ . Since  $\hat{P}(t)$  varies continuously in  $(-a, a)$ , therefore  $P_t$  does as well.  $\square$

The following convergence result is essential for proving the Trotter-type convergence formula for approximation semigroups.

**Theorem 7.8.** *Let  $\mu = \sum_{i=1}^n \frac{1}{n} \delta_{A_i} \in \mathcal{P}^1(\mathbb{P})$ . For  $\rho > 0$  let  $F_\rho := J_{\rho/n}^{\delta_{A_n}} \circ \dots \circ J_{\rho/n}^{\delta_{A_1}}$  where  $J_\rho^{\delta_A}(X) := X \#_{\frac{\rho}{\rho+1}} A$  in the spirit of (20). In particular  $F_\rho : \mathbb{P} \mapsto \mathbb{P}$  is a contraction with respect to  $d_\infty$ . For  $\lambda > 0$  let  $J_{\lambda, \rho}$  denote the approximating resolvent corresponding to  $F_\rho$  defined in (32). Then*

$$(73) \quad J_\lambda^\mu(X) = \lim_{\rho \rightarrow 0+} J_{\lambda, \rho}(X)$$

in norm, where  $J_\lambda^\mu$  is defined by (20).

*Proof.* The proof in principle is similar to the proof of Theorem 7.7 in the sense that we will use the implicit function theorem in the same way. Fix an  $X \in \mathbb{P}$  and let  $\nu := \mu + \frac{1}{\lambda} \delta_X$ . Consider the function  $F : \mathbb{R} \times \mathbb{P} \mapsto \mathbb{S}$  defined as

$$(74) \quad F(\rho, Y) := \begin{cases} \frac{1}{\rho} [J_{\rho/\lambda}^{\delta_X} \circ F_\rho(Y) - Y], & \text{if } \rho \neq 0, \\ \int_{\mathbb{P}} \log_Y \text{Ad} \nu(A), & \text{if } \rho = 0. \end{cases}$$

Notice that for  $\rho \neq 0$  we have

$$(75) \quad \begin{aligned} F(\rho, Y) &= \frac{1}{\rho} \left\{ \left( \dots (Y \#_{\frac{\rho}{\rho+n}} A_n) \#_{\frac{\rho}{\rho+n}} \dots \right) \#_{\frac{\rho}{\rho+n}} A_1 \right\} \#_{\frac{\rho}{\rho+\lambda}} X - Y \Big\} \\ &= \frac{1}{\rho} [Y_n \#_{\frac{\rho}{\rho+\lambda}} X - Y_n] + \sum_{i=0}^{n-1} \frac{1}{\rho} [Y_i \#_{\frac{\rho}{\rho+n}} A_{n-i} - Y_i] \end{aligned}$$

where  $Y_0 := Y$  and  $Y_{i+1} := Y_i \#_{\frac{\rho}{\rho+n}} A_{n-i}$  for  $0 \leq i \leq n-1$ . Thus as  $\rho \rightarrow 0$  we have  $Y_i \rightarrow Y$  and by similar calculations on each summand in the above as in the proof of Lemma 7.6, we get that

$$\lim_{\rho \rightarrow 0} F(\rho, Y) = \frac{1}{\lambda} \log_Y X + \sum_{i=1}^n \frac{1}{n} \log_Y A_i$$

in the norm topology. The convergence of the Fréchet derivative with respect to  $Y$  is a bit more delicate calculation starting from (75) using induction to calculate the derivatives of the  $Y_i$  defined recursively by composition of geometric means  $\#_{\frac{\rho}{\rho+n}}$ , but the principles are the same as in the proof of Lemma 7.6 and the calculation is left to the reader.

Now the remaining part of the proof follows the lines of the corresponding part of the proof of Theorem 7.7.  $\square$

## 8. A CONTINUOUS-TIME LAW OF LARGE NUMBERS FOR $\Lambda$

Here we combine the results of the previous sections to obtain convergence theorems valid for the nonlinear semigroups solving the Cauchy problem in Theorem 4.8.

**Theorem 8.1.** *Let  $\mu \in \mathcal{P}^1(\mathbb{P})$  and let  $\{Y_i\}_{i \in \mathbb{N}}$  be a sequence of independent, identically distributed  $\mathbb{P}$ -valued random variables with law  $\mu$ . Let  $\mu_n := \sum_{i=1}^n \frac{1}{n} \delta_{Y_i} \in \mathcal{P}^1(\mathbb{P})$  denote the empirical measures. Let  $S^\mu(t)$  and  $S^{\mu_n}(t)$  denote the semigroups corresponding to the resolvents  $J_\lambda^\mu$  and  $J_\lambda^{\mu_n}$  according to Theorem 4.6 for  $t > 0$ . Then*

$$(76) \quad \lim_{n \rightarrow \infty} S^{\mu_n}(t) = S^\mu(t)$$

*uniformly in  $d_\infty$  on compact time intervals.*

Moreover let  $F_\rho^{\mu_n} := J_{\rho/n}^{\delta_{Y_n}} \circ \dots \circ J_{\rho/n}^{\delta_{Y_1}}$  where  $J_\rho^{\delta_A}(X) := X \#_{\frac{\rho}{\rho+1}} A$  in the spirit of (20). Then

$$(77) \quad \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} (F_{t/m}^{\mu_n})^m = S^\mu(t)$$

*uniformly in  $d_\infty$  on compact time intervals.*

*Proof.* Let  $X \in \mathbb{P}$ . By Proposition 2.1,  $W_1(\mu_n, \mu) \rightarrow 0$ . Then by Theorem 3.2 we have that  $J_\lambda^{\mu_n}(Z) \rightarrow J_\lambda^\mu(Z)$  in  $d_\infty$  for any  $Z \in \mathbb{P}$ . By Theorem 4.6 we have that  $S^\mu(t)X := \lim_{m \rightarrow \infty} \left(J_{t/m}^\mu\right)^m(X)$  and  $S^{\mu_n}(t)X := \lim_{m \rightarrow \infty} \left(J_{t/m}^{\mu_n}\right)^m(X)$  uniformly on compact time intervals. The estimate

$$\begin{aligned} d_\infty\left(\left(J_{t/m}^{\mu_n}\right)^m(X), \left(J_{t/m}^\mu\right)^m(X)\right) &\leq d_\infty\left(\left(J_{t/m}^{\mu_n}\right)^m(X), \left(J_{t/m}^{\mu_n}\right)^{m-1} J_{t/m}^\mu(X)\right) \\ &\quad + d_\infty\left(\left(J_{t/m}^{\mu_n}\right)^{m-1} J_{t/m}^\mu(X), \left(J_{t/m}^\mu\right)^{m-1} J_{t/m}^\mu(X)\right) \end{aligned}$$

by induction yields the existence of  $n_0 \in \mathbb{N}$ , such that for  $n > n_0$  we get that

$$d_\infty\left(\left(J_{t/m}^{\mu_n}\right)^m(X), \left(J_{t/m}^\mu\right)^m(X)\right) < \epsilon$$

for a fixed  $m$ , thus by the triangle inequality we get (76). Uniform convergence can be showed similarly along the lines of the proof of Proposition 6.9.

Now (77) follows from (76) combined with Theorem 7.8 and Theorem 6.10.  $\square$

**Acknowledgments.** The work of Y. Lim was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government(MEST) No.2015R1A3A2031159. The work of M. Pálfi was supported in part by the "Lendület" Program (LP2012-46/2012) of the Hungarian Academy of Sciences and the National Research Foundation of Korea (NRF) grant funded by the Korea government(MEST) No.2016R1C1B1011972.

## REFERENCES

- [1] L. Ambrosio, N. Gigli and G. Savaré, Gradient flows in metric spaces and in the space of probability measures, Birkhäuser Verlag, Basel, 2005.
- [2] M. Bačák, Convex analysis and optimization in Hadamard spaces, De Gruyter Series in Nonlinear Analysis and Applications 22 (2014).
- [3] R. Bhatia and J. Holbrook, Riemannian geometry and matrix geometric means, Linear Algebra Appl. 413 (2006), pp. 594–618.
- [4] R. Bhatia and R. Karandikar, Monotonicity of the matrix geometric mean, Math. Ann. 353:4 (2012), pp. 1453–1467.

- [5] M. G. Crandall and T. M. Liggett, Generation of Semi-Groups of Nonlinear Transformations on General Banach Spaces, *Amer. J. Math.*, Vol. 93, No. 2 (1971), pp. 265–298.
- [6] J. Diestel and J. J. Uhl, *Vector Measures*, American Mathematical Society, Providence, Rhode Island (1977), 322 pp.
- [7] V. I. Bogachev, *Measure Theory Vol II.*, Springer-Verlag Berlin Heidelberg (2007), 575 pp.
- [8] J. I. Fujii, Operator Means and the Relative Operator Entropy, *Operator Theory and Complex Analysis Vol. 59, Operator Theory: Advances and Applications*, pp. 161–172.
- [9] J. I. Fujii and Y. Seo, The relative operator entropy and the Karcher mean, *Linear Algebra Appl.* (2016), <http://dx.doi.org/10.1016/j.laa.2016.11.018>.
- [10] D. Gatzouras, On weak convergence of probability measures in metric spaces, unpublished manuscript (2001), preprint available: <http://users.math.uoc.gr/gatzoura/>.
- [11] J. Holbrook, No dice: a deterministic approach to the Cartan centroid, *J. Ramanujan Math. Soc.* **27:4** (2012), pp. 509–521.
- [12] H. Karcher, Riemannian center of mass and mollifier smoothing, *Comm. Pure Appl. Math.*, vol. 30 (1977), pp. 509–541.
- [13] A. B. Kharazishvili, *Set Theoretical Aspects of Real Analysis*, Monographs and Research Notes in Mathematics, CRC Press, Taylor & Francis Group 6000 Broken Sound Parkway NW (2015), 452 pp.
- [14] S. Kim and H. Lee, The power mean and the least squares mean of probability measures on the space of positive definite matrices, *Linear Algebra Appl.* 465 (2015), pp. 325–346.
- [15] F. Kubo and T. Ando, Means of positive linear operators, *Math. Ann.*, 246 (1980), pp. 205–224.
- [16] J. Lawson, Ordered Probability Spaces, preprint (2016), arXiv:1612.03213.
- [17] J. Lawson and Y. Lim, Symmetric spaces with convex metrics, *Forum Math.* **19** (2007), pp. 571–602.
- [18] J. Lawson and Y. Lim, Karcher means and Karcher equations of positive definite operators, *Tran. of the AMS*, 1 (2014), pp. 1–22.
- [19] Y. Lim and M. Pálfi, Matrix power means and the Karcher mean, *J. Func. Anal.*, Vol. 262, No. 4 (2012), pp. 1498–1514.
- [20] Y. Lim and M. Pálfi, Weighted deterministic walks for the least squares mean on Hadamard spaces, *Bull. of the London Math. Soc.* 46:3 (2014), pp. 561–570.
- [21] L. H. Loomis and S. Sternberg, *Advanced Calculus*, Jones and Bartlett Publishers, London, 1990.
- [22] I. Miyadera and S. Oharu, Approximation of semi-groups of nonlinear operators, *Tôhoku Math. J.* 22 (1970), pp. 24–47.
- [23] M. Moakher, A Differential Geometric Approach to the Geometric Mean of Symmetric Positive-Definite Matrices, *SIAM J. Matrix Anal. Appl.*, 26 (2005), pp. 735–747.
- [24] A. Navas, An  $L^1$  ergodic theorem with values in a nonpositively curved space via a canonical barycenter map, *Ergodic Theory and Dynamical Systems* 33 (2013), pp. 609–623.
- [25] S. Ohta and M. Pálfi, Discrete-time gradient flows and law of large numbers in Alexandrov spaces, *Calc. Var. PDE* 54 (2015), pp. 1591–1610.
- [26] S. Ohta and M. Pálfi, Gradient flows and a Trotter-Kato formula of semi-convex functions on CAT(1)-spaces, to appear in *Amer. J. Math.* (2016).
- [27] M. Pálfi, Operator means of probability measures and generalized Karcher equations, *Adv. Math.* 289 (2016), pp. 951–1007.
- [28] I. Stojkovic, Approximation for convex functionals on non-positively curved spaces and the Trotter-Kato product formula, *Adv. Calc. Var.* 5 (2012), pp. 77–126.
- [29] K.-T. Sturm, Nonlinear martingale theory for processes with values in metric spaces of non-positive curvature, *Ann. Probab.* **30** (2002), pp. 1195–1222.
- [30] M. Takesaki, *Theory of Operator Algebras Vol 1*, Springer-Verlag, Berlin, 2002.
- [31] A.C. Thompson, On certain contraction mappings in a partially ordered vector space, *Proc. Amer. Math. Soc.*, 14 (1963), pp. 438–443.
- [32] C. Villani, *Topics in Optimal Transportation (Graduate Studies in Mathematics, 58)*, Springer, Berlin, 2003.

DEPARTMENT OF MATHEMATICS, SUNGKYUNKWAN UNIVERSITY, SUWON 440-746, KOREA AND  
MTA-DE "LENDÜLET" FUNCTIONAL ANALYSIS RESEARCH GROUP, HUNGARIAN ACADEMY OF SCI-  
ENCES, 1051 BUDAPEST, SZÉCHENYI ISTVÁN SQR. 9, HUNGARY.

*E-mail address:* `palfia.miklos@aut.bme.hu`